Plasticity of the Human Auditory Cortex and of its Connections to Other Modalities

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(**Dr. sc. nat.**)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Simon Baumann

von

Grindelwald BE

Promotionskomitee

Prof. Dr. Kevan Martin (Vorsitz)

Prof. Dr. Lutz Jäncke (Leitung der Dissertation)

Prof. Dr. Stephan Neuhauss

Dr. Martin Meyer

Zürich 2006

Statement of Contributions

Contributing authors:

Simon Baumann (SB)

Martin Meyer (MM)
Lutz Jäncke (LJ)
Susan Koeneke (SK)
Kai Lutz (KL)
Conny F. Schmidt (CFS)
Sarah Marchina (SM)

Details of contribution to studies:

I

- Idea	SB
- Concept/design of study	SB, MM
- Design of stim. material/ present. program	SB
- Organization of Subjects	MM, SB
- Data recording	2 students u.s.o.[1] MM, SB
- Data analysis	SB
- Data interpretation	SB, MM
- Manuscript	MM w.c.o.[2] SB

II

- Idea	LJ
- Concept/design of study	MM, SB
- Design of stim. material/ present. program	SB
- Organization of Subjects	2 students u.s.o. MM, SB
- Data recording	SB, 2 students u.s.o. SB
- Data analysis	SB
- Data interpretation	SB, MM
- Manuscript	SB w.c.o. MM

III

- Idea	SB
- Concept/design of study	SB, SK, KL
- Design of stim. material/ present. program	SB, KL
- Organization of Subjects	SB, SK
- Data recording	SB, KL, SK w.a.o.[3] CFS
- Data analysis	SB, SK, KL
- Data interpretation	SB, SK, KL
- Manuscript	SB w.c.o. SK, KL, MM

[1] under supervision of
[2] with contribution of
[3] with assistance of

IV

- Idea	LJ, MM
- Concept/design of study	MM w.c.o. SB, SM
- Design of stim. material/ present. program	SB, SM
- Organization of Subjects	SM
- Data recording	SB w.a.o. SM, CSF
- Data analysis	SM u.s.o. SB, MM
- Data interpretation	MM, SB, LJ
- Manuscript	MM w.c.o. SB

All studies were financed by a SFN grant to LJ. He had the overall supervision, approved the experiments and made the final decision on the manuscripts.

Figures are designed by SB.

Table of Contents

1 Zusammenfassung 7
2 Summary 9
3 Introduction 11
3.1 A brief introduction into the organization of the auditory cortex 13
3.2 Introduction into plasticity 16
3.2.1 Plasticity in the auditory cortex of animals and humans 18
3.2.2 Music and plasticity 19
3.3 Motivation for the individual studies 21
4 First Study 24
4.1 Abstract 24
4.2 Introduction 26
4.3 The present study 30
4.4 Materials and Methods 31
4.5 Results 37
4.6 Discussion 45
4.7 Conclusion 53
4.8 References 54
5 Second Study 61
5.1 Abstract 61
5.2 Introduction 63
5.3 Material and Methods 64
5.4 Results 69
5.5 Discussion 76
5.6 Conclusion 80
5.7 References 80
6 Third Study 83
6.1 Abstract 83
6.2 Introduction 85
6.3 Materials and Methods 88
6.4 Results 93
6.5 Discussion 100
6.6 References 106
7 Fourth Study 112
7.1 Abstract 112
7.2 Introduction 113
7.3 Materials and Methods 117
7.4 Results 122
7.5 Discussion 128
7.6 Conclusion 133
7.7 References 133
8 Discussion, Conclusion & Outlook 139
8.1 A brief summary of the results 139
8.2 Are the extraordinary abilities of musicians acquired or innate? 140
8.3 Expertise and increased neuronal representations of stimuli 142
8.4 Top down and crossmodal influences on the auditory cortex 144
8.5 Conclusions & Outlook 146
8.6 References 148

1 Zusammenfassung

Der auditorische Kortex spielt eine entscheidende Rolle für höhere kognitive Funktionen wie z. B. das Verstehen von Sprache, Musik und auditorischem Raum. Das reibungslose funktionieren der auditorischen Modalität hängt von der Fähigkeit seiner kortikalen Strukturen ab, sich den ändernden Herausforderungen unserer Umwelt anzupassen. Diese Anpassungsfähigkeit der kortikalen Architektur durch Modifikation ihrer strukturellen und funktionalen Eigenschaften nennt man Plastizität. Ohne die Annahme eines plastischen auditorischen Kortex und plastischen Verbindungen zu anderen Gehirnarealen wären wir nicht in der Lage zu erklären, wie es einem neugeborenen Kind möglich ist eine Sprache zu erlernen. Es wäre schwierig zu verstehen, wie eine erwachsene Person sich von einer schweren Hirnschädigung oder einem Infarkt erholt bei dem sie ihre Fähigkeit zu Sprechen verloren hat. Obwohl erste Hinweise aus Studien in Tieren und Menschen darauf hindeuten, dass sich der auditorische Kortex durch Training in seinen funktionellen und strukturellen Eigenschaften ändert, ist noch wenig über die dafür notwendigen Voraussetzungen und über die Grenzen bekannt. Die vier in dieser Dissertation enthaltenen Studien dienen dazu, diese Fragen anzugehen. In den ersten beiden Studien wurde das neuronale, durch Instrumentaltöne und Sinustöne ausgelöste, elektrische Potential in Musikern analysiert und mit Nichmusikern verglichen um Veränderungen zu suchen, welche auf das intensive Training der Musiker hinweisen. Im Unterschied zu vorhergehenden Studien fanden wir nicht nur ein erhöhtes auditorisches Potential für die geübten Instrumentaltöne in Musikern verglichen mit Nichtmusikern sondern auch für die Sinustöne. Dieses Resultat deutet darauf hin, dass Musiktraining nicht nur zur Veränderung der neuronalen Antwort auf geübte Stimuli führt, sondern dass Training die auditorische Verarbeitung im Allgemeinen verändert und möglichweise verbessert. Ausserdem konnten wir zeigen, dass die erwähnte Veränderung nicht durch eine vorübergehende Aufmerksamkeits-verlagerung zustande kam, was die Hypthese einer dauerhaften Veränderung im auditorischen Kortex durch Training erhärtet.

In der dritten und vierten Studie untersuchten wir mittels funktioneller Magnetresonanz (fMRI) die Interaktion des auditorischen Kortex mit der motorischen und dem visuellen Modalität. Unsere Daten zeigen teilweise überlappende Netzwerke kortikaler Areale auf, welche eine wichtige Rolle beim audio-motorischen und audio-visuellen Informationstransfer spielen. Das gefundene Netzwerk ist durch eine zunehmende Anzahl an Studien anderer Gruppen bestätigt worden. Die Daten von Musikern der dritten Studie deuten klar darauf hin, dass die transmodale Aktivität durch intensives Piano Training zunimmt. Schlussendlich fanden wir in der vierten Studie, nach einem kurzen Durchgang von Paarungen auditorischer und visueller Reize, Hinweise für eine erhöhte auditorische Aktivität in Regionen des, aus den vorangehenden Studien abgeleiteten, audio-visuellen Netzwerkts. Die Aktivität im sekundären auditorischen Kortex wurde einzig durch die Präsentation der visuellen Stimuli ausgelöst. Dieses Resultat deutet darauf, dass audio-visuelle Verbindungen durch Training verstärkt werden können.

Zusammenfassend bestätigen die in dieser Dissertation beschriebenen Resultate die Existenz des Phänomens Plastizität im auditorischen Kortex und sie fügen wertvolle Informationen über die dem Phänomen zugrunde liegenden Mechanismen hinzu. Im Besonderen heben diese neuen Daten die Rolle von transmodalen Interaktionen für die Plastizität im auditorischen Kortex hervor. Aus diesem Grund ist diese Forschungarbeit ein Fortschritt für das Verständnis der Plastizität im auditorischen Kortex und sie bildet eine optimale Grundlage für weitere Untersuchungen der transmodalen Integration.

2 Summary

The auditory cortex has a crucial role in higher cognitive functions including the perception of speech, music and auditory space. The proper functioning of the auditory modality depends on the capacity of its cortical circuits to adapt to the changing demands posed by our environment. The capacity of cortical structures to adapt by changing their structural and functional properties is called plasticity. Without assuming a plastic auditory cortex and plastic connections to other cortical areas we would not be able to explain how a newborn child is able to learn speech and it would be difficult to understand how an adult person recovers after losing their ability for speech by brain injury or stroke. Although first evidence from animals and humans suggests modification of the auditory cortex in its functional and structural properties due to training, little is known about the preconditions and limits of these changes. In order to address these questions the four studies included in this thesis were conducted. In the first two studies the neural electric potentials evoked by instrumental tones and sine wave tones were analyzed in musicians and compared to nonmusicians in order to search for changes in the neural response due to intensive music training. In contrast to previous studies we observed enhanced auditory potentials in musicians compared to nonmusicians not only for the trained instrumental tones but also for sine wave tones. This result suggests that in addition to the modification of neural responses to trained stimuli music training changes and possibly improves auditory processing in general. Furthermore, we were able to demonstrate that the reported change was not due to a transient attention shift and thus implies a permanent alteration of the auditory cortex by training.

In the third and fourth study we investigated the interaction of the auditory cortex with the motor- and the visual modality using functional Magnetic Resonance Imaging (fMRI). Our data outlined partly overlapping networks of cortical areas that play an important role in audio-motor and audio-visual information transfer. The proposed network has been confirmed by an increasing body of data from other groups. The data of the third study derived from professional pianists clearly indicate that

crossmodal activity in the audio-motor network is increased due to intensive piano training. Finally, in the fourth study of this thesis, we observed evidence for increased auditory activity in areas of the proposed audio-visual network triggered by visual stimuli alone after a short pairing session of unrelated auditory and visual stimuli. These results suggest that training enhances audio-visual connectivity.

In summary, the results presented in this thesis confirm the existence of phenomena of plasticity in the human auditory cortex and they add valuable information on possible underlying mechanisms. In particular, new data reported highlighting the role of crossmodal interactions for plasticity in the auditory cortex. Thus, this piece of research is a step forward for the understanding of auditory plasticity and it forms an optimal basis for further investigations of crossmodal integration.

3 Introduction

Audition is one of our most important senses. Vital human abilities depend on the perfect functioning of the auditory modality. Not only typically human cognitive functions such as perception of speech and music would be impossible without a working auditory system but we would lack the ability to constantly monitor what happens in our immediate environment outside of our visual field. Orienting in a world of sounds would be massively complicated. Nevertheless, we often underestimate the importance of the auditory system for our daily life and our conscious perception is completely dominated by vision. This might be one of the reasons why much less research has been conducted investigating the auditory cortex compared to other cortical sensory areas e.g. the visual cortex. It could also be an explanation why research on the auditory cortex was much slower to recognize the relevance of cortical self-organization and reorganization, the most obvious elements of cortical plasticity.

Plasticity is an intrinsic property of the human brain and describes the fact that its structure is not completely predefined by our genome. In contrast, for instance, to the computers we use every day, the brain does not have a rigid architecture but it is a highly dynamic network that once established at the earliest stage of our development, is constantly modified during our whole life-time. Plasticity is the precondition for our capacity to adapt to the changing requirements posed by our natural and social environment ensuring that we can learn from our experiences and the experiences of others. A better understanding of our brain's potential for plasticity does not only highlight the limits of our capacity to learn but it could also reveal ways to extend these limits. Furthermore, profound knowledge of the mechanisms of plasticity are crucial to understand how our brain recovers from injuries of the cortex, brain diseases or stroke so we can actively support the healing processes.

Plasticity in the auditory cortex is critical for the understanding of several important functions of our brain. A plastic brain is, for example, of particular relevance to the acquisition of language. Even with an innate capacity for language, normal speech

can hardly be acquired without auditory feedback[4] and a capacity for learning. It is therefore no surprise that many speech disorders have been associated with plasticity deficits (Kraus et al., 1996; Merzenich et al., 1996; Tallal et al., 1996). Likewise, a system capable of localizing sound with extraordinary precision, using various sets of cues, cannot be achieved without resorting to tuning mechanisms that recalibrate the system continually, especially during the growth phase of the head and outer ears. Acoustic communication and spatial processing are only two of many examples to illustrate the importance of cortical plasticity in the auditory system for many cognitive functions we use every day.

The following three sections provide an introduction into the organization of the auditory cortex and a brief overview on the state of research in auditory plasticity. Recent research is presented that has observed first evidence for plasticity occurring in the auditory cortex. However, it is also highlighted that the field is still at the beginning to conceive under which circumstances plasticity is possible and it is far away from understanding the exact mechanisms underlying the modification of the auditory cortex. The principal aim of this thesis is therefore to address some of the open questions in order to reduce the gaps in our knowledge on the functioning of the auditory cortex and to increase our understanding in the complex mechanisms that finally lead to plasticity. At the end of this introduction the two main questions motivating this thesis are formulated and the precise objectives of the experiments are outlined.

[4] That the same audio-motor feedback circuits are also crucial for instrumental music performance is demonstrated in the third study of this thesis.

3.1 A brief introduction into the organization of the auditory cortex

The functional anatomy of the auditory cortex has not been as thoroughly investigated as the visual cortex, at all. Nevertheless, considerable knowledge has been acquired particularly on the anatomy and the cytoarchitecture of the auditory cortex. Most of the functional data is derived from monkeys. Only the emergence of modern imaging technologies at the end of the last century permitted a systematic investigation of the functional organization of the auditory cortex in humans. Meanwhile, it is undisputed that in both monkeys and humans most of the basic auditory functions are processed in the superior temporal gyrus (STG) and the supra temporal plane (STP) buried deep in the Sylvian fissure (See also Fig. 1).

Electrophysiologists proposed a model for the organization of the auditory cortex of macaques which comprises core, belt and parabelt areas (Fig. 1) (reviewed in Kaas and Hackett, 1998; Rauschecker, 1998). The core area is located in the center of the auditory cortex and has a nearly anterior to posterior orientation. The auditory core is further divided into three subfields each of which seems to be organized in a tonotopic manner. The core areas receive ascending inputs from the ventral medial geniculate body and project to ipsilateral and contralateral core areas, as well as to adjacent belt areas. Neurons in the core respond well and with short latencies to pure tones, with narrow frequency tuning at their characteristic frequency (Rauschecker et al., 1995; 1997). The surrounding belt region includes 7 or 8 subfields which are not considered to be primary areas. Each belt field receives major inputs from the adjacent core field and from the dorsal and medial divisions of the medial geniculate body. Neurons in the belt area have more complex receptive field properties than those in the core, probably reflecting integration over convergent inputs. For example, in the lateral belt region, neurons respond vigorously to spectrally complex stimuli such as vocalizations (Rauschecker et al., 1997). The physiological properties of the parabelt area are less established. This area receives afferents largely from neurons in the belt. Projections from the parabelt seem to end mainly in various regions of the temporal cortex, but there are also connections to the parietal lobe and the prefrontal cortex reported (Romanski et al., 1999).

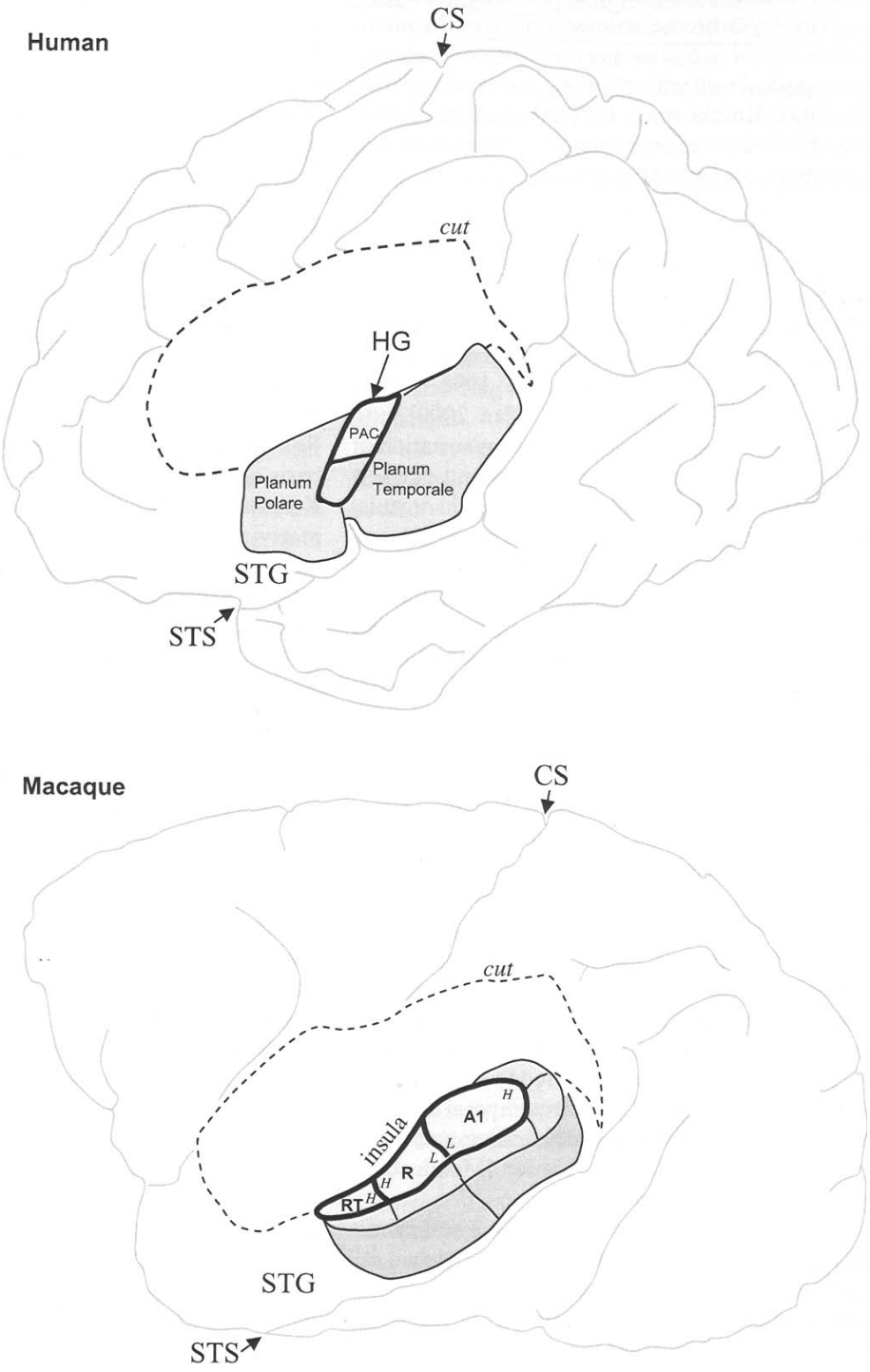


Fig. 1. *Auditory cortex in human and macaque.* Dorsolateral view of the human and macaque cerebral cortex after removal of the overlying parietal cortex. This view exposes the superior surface of the superior temporal gyrus (STG). The dashed line defines the portion of the cortex that has been cut away. The primary auditory cortex (PAC) in humans and its corresponding areas in macaques is depicted in light grey. Belt and parabelt areas are shown in darker grey. Subfields (A1, R, RT) within the auditory core are denoted and high-frequency (H) and low-frequency response regions indicate the orientation of the tonotopic gradients in these core fields. CS = Central Sulcus; STS = superior temporal sulcus. Image by Hall et al. (2003).

Although the functional architecture of the auditory cortex in humans is much less established compared to monkeys, humans seem to share the hierarchical organization described above and known from various animals. The anatomy of the human STP is typically governed by three large cortical formations (see also Fig. 1). The Heschl's Gyrus (HG), the first transverse gyrus, is located approximately in the center of the STP and has an orientation diagonal to the anterior-posterior and medial-lateral axis. The medial two thirds are considered to house the primary auditory cortex and are therefore the equivalent of the auditory core areas in monkeys. The primary auditory cortex in humans is believed to be organized in simple tonotopic manner showing a gradient of responding neurons from low frequencies at the anterolateral side to high frequencies at the posterior-medial side. Recent high resolution imaging data, however, presents a more complex situation (see e.g. Formisano et al., 2003). It is not yet resolved whether methodological limits are responsible for this result or whether tonotopy is a more complex phenomenon than originally believed.

The Planum Temporale (PT) covers the posterior part of the temporal lobe, adjacent to the HG. The Planum Polare (PP) is located on the anterior side of the HG. The PT and the PP are considered to be equivalents of the monkey belt areas. However, systematical comparisons between the functions of monkey and human belt areas are sparse. At least, recent data on the cytoarchitecture of the human auditory cortex revealed a surprising convergence of macroscopic landmarks such as HG, PT and PP with areas divided by cytoarchitectural measures (Morosan et al., 2001; Morosan et al. personal communication). That is supporting evidence for the functional meaningfulness of these macroscopic landmarks, which are widely used in the field of neuroimaging to interpret the recorded functional data. A third macroscopic area that is traditionally counted to the auditory cortex is the lateral convexity of the STG. From its anatomical location and its cytoarchitectonic properties, this area is the best cadidate for an equivalent to the monkey lateral belt area. However, systematic analysis of this area is sparse in either species.

Although a considerable number of neuroimaging studies have been accomplished to establish the functional properties of the human auditory cortex, no clear and simple partitioning into areas and subareas with distinct and well defined functional properties has emerged up to now. Contradictory results have been observed for many issues and the functional organization has been investigated on too many

levels (from physical properties up to the convergence with complex psychological, linguistical and even philosophical theories) to draw a simple conclusion. A thorough discussion of all these results would exceed by far the scope of this brief introduction. Hence, only the leas disputable results are reviewed in this section.

In agreement with the results from monkeys the human primary cortex (AI) may have a larger involvement in processing basic sound features, such as frequency and level. Whereas pure tones seem to evoke similar levels of activity in AI as in the surrounding (mainly posterior) secondary areas such as the PT, the latter areas seem to play a more important role in the processing of spectrotemporally complex sounds. In addition, increasing evidence (see also the third and fourth study of this thesis) suggests that the area posterior to the AI plays an important role in the interaction of the auditory cortex with other modalities. The areas anterior of the HG including the PP are the least investigated. A personal interpretation of the existing data leads to the conclusion that particularly complex stimuli with modulation of sound features on a timescale of seconds such as language and melodies show strong activity in these areas. As a matter of course, many more areas in the brain are activated by sound and complex auditory stimuli. But only the areas described in this section are indisputably involved in the specific processing of auditory information.

3.2 Introduction into plasticity

Already at the beginning of the last century the famous neuroanatomist and pioneer in neuroscience Santiago Ramón y Cajal understood that the acquisition of new skills can only be explained by the modification of the neuronal network either by reinforcing pre-established organic pathways or by the formation of new connections. After the neuron was recognized by Cajal and his colleagues Auguste Forel and Wilhelm His to form the building block of the nervous system, generations of physiologists and molecular biologists have investigated the electro-chemical properties and molecular mechanisms of single neurons and small neuronal networks. This research resulted in theories such as “Hebb’s law” by Donald Hebb (1949) proposing that coincident activity in two connected neurons would lead to

strengthening of their connection. This theory allowed Hebb to explain a least some forms of associative learning by simple neuronal mechanisms. The basic idea of “Hebb’s Law“ was later confirmed by the discovery of “Long-term potentiation” (LTP) in the dentate gyrus of the hippocampal formation. LTP is defined as a long –lasting increase of synaptic efficacy following a brief high-frequency stimulation and it was discovered by Bliss & Lomo (1973). It was seen as a promising mechanism for learning and memory and further research could explain the detailed mechanisms on a molecular level (reviewed in Chen and Tonegawa, 1997). Whereas in the years following the discovery of LTP extensive research has been conducted to increase the understanding of the hippocampus’ role for learning and memory, the sensory systems have traditionally been viewed as “stimulus analyzers” with little potential for plasticity. Although visual deprivation studies in cats (reviewed in Sherman and Spear, 1982) and similar studies in the somatosensory cortex (Kaas, 1983) clearly demonstrated the potential for plasticity in the visual and somatosensory cortices, the reported changes mainly occurred within a “critical period” of development. However, it was shown in later experiments that sensory cortices of various modalities were also susceptible to plastic alterations in adult animals by manipulation of the sensory input. The somatotopic representation of fingers and limbs in primary sensory cortex SI has made this modality particularly suitable for the investigation of plasticity because the input manipulation resulted in considerable map reorganization. In New World monkeys it was shown for example that the deactivation of inputs from half of the hand, which deprived a large lateral sector of the hand representation, resulted in finger representations which were many times larger than in normal animals (Merzenich et al., 1983). Interestingly, the effects of deactivating a peripheral nerve have been partly or totally reversed after the regeneration of the nerve. Yet, not only destructive manipulation has proved to influence the cortex of higher animals but Jenkins et al. (1990) demonstrated that also training had a significant effect on sensory areas. In their study monkeys were trained to maintain contact with a rotating disk with one or two digits which resulted in expansion of the cortical representation of these digits.

In the late eighties a large number of studies showed similar map reorganization phenomena after deafferencation, reafferencation and training in the somatosensory, the visual and the motor cortex of adult animals (reviewed in Kaas, 1991). This convincing amount of data has led to a revision of the traditional view of the sensory

cortices as static structures that show little change once established. Research on plasticity in the auditory cortex, however, was not widespread in those days and evidences for plasticity in the cortical areas of the auditory system mainly came up at the beginning of the nineties.

3.2.1 Plasticity in the auditory cortex of animals and humans

Robertson et al. (1989) could demonstrate that restricted lesions in the tonotopically organized cochlea resulted in plasticity of the cochleotopic frequency map in the primary auditory (AI). AI neurons deprived of frequency input they were normally most sensitive to have acquired sensitivity to frequencies represented at cochlear sites adjacent to the lesion. Later, the same group could demonstrate that the observed effects reflected plastic changes rather than simply being the residue of pre-lesion input (Rajan et al., 1993). Recanzone et al. (1993) showed later that training could influence the map organization of the auditory cortex, too. In AI of adult owl monkeys they reported an increase in size of a cortical area representing a restricted frequency range that was correlated with the animal's performance discrimination task for this specific frequency. The monkeys were trained for several weeks to discriminate small differences in frequency of sequentially presented tones. Even simple forms of associative learning were shown relatively early to modify the auditory cortex. Classical conditioning has specifically modified receptive fields in the AI and secondary auditory cortical areas that they favored the frequency of a tone signal over other frequencies (reviewed in Weinberger, 1993). Indirect evidence for plasticity was derived by more theoretical considerations based on observations in guinea pigs. It turned out that the large variability in the shapes and sizes of the guinea pigs' external ears were matched by their unique "location specific attenuations" or "head-related transfer functions" (HRTFs) (Sterbing et al., 1996). The authors claimed that the complex interactions between ear shape and HRTFs could hardly be anticipated by genetic mechanisms and thus are an indication of plastic adaptation.

Auditory plasticity in humans is either reported from the research field of deafness and cochlear implants or from music research. Whereas several studies could

demonstrate a modification of responses after acquired deafness or by stimulation of neurons using cochlear implants they suffered from the same drawbacks as animal deprivation studies that the functionality of these modifications remained controversial (reviewed in Middlebrooks et al., 2005). An implementation of training studies comparable to the studies accomplished in animals that could successfully demonstrate auditory plasticity is not simple because it is difficult to find human participants who are ready to train e.g. simple tone discrimination tasks for weeks or months in a controlled laboratory environment. The ideal alternative would be to have persons who already experienced a relatively standardized auditory training not only for months but even for years. That is the point where musicians become interesting.

3.2.2 Music and plasticity

In the recent years, music became a popular topic in neuroscience. The popularity of music in brain research has many reasons including the fact that music is also very popular outside neuroscientific laboratories. For researchers interested in plasticity, however the intensive training and the extraordinary manual and perceptual abilities of musicians are of particular interest because it makes musicians an ideal model system to study what changes in the brain when exercise is turned into ability. Interestingly, one of the pioneers in neuroscience has already used the extraordinary training and skill of pianists to explain the necessity of a plastic nervous system. The following text passage is an excerpt from the well known publication *Textura del Sistema Nervioso* of the year 1904 by Santiago Ramón y Cajal:

> La labor de un pianista […] es inaccesible para el hombre ineducado ya que la adquisición de nuevas habilidades requiere muchos años de práctica mental y física. Para entender plenamente este complejo fenómeno se hace necesario admitir, además del refuerzo de vias orgánicas pre-establecidas, la formación de vias nuevas por

> ramificación y crecimiento progresivo de la arborización dendrítica y terminales nerviosas.[5] (p. 296)

In particular, the intensity of the auditory training experienced by musicians is arguably unparalleled by any other profession or activity. Ericson et al. (1993) calculated that the accumulated practice time for the best violin students at the Berlin Academy of Music was 7400 h by the time they were 18 years old neglecting the time they were simple listening to music. This is in contrast to 3400 h for violinists who were student music teachers and to 1600 h for amateur pianists. Indeed, the intensive training and/or the early commencement of training seem to have an impact on macroscopic brain structures. Schneider et al. (2002) reported an increased size of the Heschl's Gyrus in musicians. Gross anatomical peculiarities were also observed in the motor cortex (Amunts et al., 1997; Gaser and Schlaug, 2003) and the Corpus Callosum (Schlaug et al., 1995) of musicians.

Functional evidence for plasticity in musicians was mainly observed in Electroencephalography (EEG) and Magnetoencephalography (MEG) studies. A seminal study by Pantev et al. (1998) reported an increased auditory evoked field for piano tones compared to sine wave tones in pianists. A similar difference was not observed for nonmusicians in this study. The auditory evoked field (AEF) is the magnetic component of the neuronal response to auditory events. The particular time range in which Pantev observed the field difference is called Negativity 1 (N1), represents an early cortical response and is widely accepted to be generated in the primary or secondary auditory cortex. Thus, the particular role of piano tones in the AEF of pianists was interpreted as an enhanced response elicited by an enlarged neuronal representation for piano tones due to intensive piano training, referring to the results of the tone training study by Recanzone et al. (1993). This interpretation suggested plastic reorganization of the auditory cortex due to long term exposure to specific sound. Several following studies reported various components in musicians

[5] The labor of a pianist […] is inaccessible for the uneducated man as the acquisition of new skill requires many years of mental and physical practice. In order to fully understand this complex phenomenon it becomes necessary to admit, in addition to the reinforcement of pre-established organic pathways, the formation of new pathways through ramification and progressive growth of the dendritic arborization and the nervous terminals. (translation by Alvaro Pacual-Leone, 2005)

that showed an increased response to instrumental tones (Kuriki et al., 2006; Shahin et al., 2003; 2005).

The data presented in this section shows first evidence for plasticity in auditory cortex of animals and humans. In particular, the results from the training studies and from the studies with musicians suggest that the observed cortical alterations are indications of improved functionality. These encouraging results form the background that motivated several lines of studies in order to improve our understanding of auditory plasticity and to address some of the many unresolved questions. Four of these studies form the basis of this thesis. Whereas the precise objectives of each study are outlined in the following paragraphs all of the four works are pieces in the puzzle to answer the following two questions:

Does an intensive auditory training result in a permanent modification of activity in the human auditory cortex?

Does the training of crossmodal interactions result in increased connectivity of the auditory cortex with other modalities?

3.3 Motivation for the individual studies

First study: EEG pilot data sets recorded from nonmusicians during the preparation of a study to investigate the auditory evoked potential in musicians and nonmusicians showed evidence for a systematically enhanced AEP for instrumental tones compared to intensity matched sine wave tones. These findings were contradictory to the results reported by Pantev et al. (1998) and a confirmation would have weakened the argumentation that an enhanced AEP for piano tones compared to sine wave tones in pianists indicates plasticity. Furthermore, systematically increased neural responses to complex instrumental tones compared to sine wave tones suggests the recruiting of additional neural resources possibly activated by the additional harmonics present in instrumental tones but not in sine wave tones. Cortical structures specifically responding to acoustic harmonics may provide insight in the

processing of timbre, an important sound feature that was hardly investigated at this time. Thus, the main objectives of this study were thus: 1. The confirmation of an increased response of instrumental tones compared to sine wave tones in nonmusicians. 2. The localization of the cortical structures which are responsible for this difference.

Second Study: The finding from the first study that both musicians and nonmusicians differed in the response to instrumental tones compared to sine wave tones put the original interpretation of a long-term exposure to piano tones indicating plasticity into perspective. Furthermore, a transient increase of the AEP in musicians due to attention had to be excluded before proposing a permanent cortical modification. The main objectives of this study were therefore: 1. The direct comparison of the AEPs of musicians to nonmusicians. 2. A systematic investigation of the influence of attention to potential differences in the AEP magnitude between musicians and nonmusicians.

Third Study: The acquisition of new skills does not only require a modification of the sensory and motor cortices. New connections between modalities or the modification of preexisting connection are crucial for many tasks. For instance, a pianist has to coordinate the auditory input with the motor output by feedforward and feedback mechanisms to ensure a perfect artistic outcome. A previous EEG study (Bangert et al., 2001) reported similar scalp potential maps for music playing and music listening in professional pianists whereas the same motor and auditory maps in piano novices matched the better the longer they were training piano performance. The authors proposed a common network of interacting auditory and motor areas that was established due to intensive piano training. The main objectives of this study were: 1. The identification of cortical structures that play an important role in audio-motor interaction. 2. to test whether we are able to find increased interactions of auditory and motor areas in musicians compared to nonmusicians as an indication of plasticity due to long-term audio-motor interaction training.

Fourth study: Previous studies (Bunzeck et al., 2005; Jancke and Shah, 2004) demonstrated that imagery of sounds triggered by related pictures could evoke activity in the secondary auditory cortex. In order to investigate short term plasticity of audio-visual interactions the objective of this study was to investigate whether a repetitive pairing of unrelated visual and auditory stimuli would finally result in similar auditory activity triggered by the visual stimulus alone.

4 First Study

Electrical Brain Imaging Reveals Spatio-Temporal Dynamics of Timbre Perception in Humans

MARTIN MEYER, SIMON BAUMANN & LUTZ JÄNCKE

Div. Neuropsychology, Institute for Psychology, University of Zurich

Published in *Neuroimage (2006). 32, 1510-23*

4.1 Abstract

Timbre is a major attribute of sound perception and a key feature for the identification of sound quality. Here we present event-related brain potentials (ERPs) obtained from sixteen healthy individuals while they discriminated instrumental tones (piano, trumpet, and violin) or simple sine wave tones that lack the principal features of timbre. Data analysis yielded enhanced N1 and P2 responses to complex instrumental tones relative to sine wave tones. Furthermore, we applied an electrical brain imaging approach using low-resolution electromagnetic tomography (LORETA) to estimate the neural sources of N1/P2 responses. Significance tests of instrumental vs. sine wave tones for N1 and P2 separately revealed distinct regions as principally governing timbre perception. In an initial stage (N1), timbre perception recruits left and right (peri-)auditory fields with an activity maximum in the right posterior Sylvian fissure (SF) and the posterior cingulate (PCC) territory. In the subsequent stage (P2) we uncovered enhanced activity in the vicinity of the entire cingulate gyrus. The involvement of extra-auditory areas in timbre perception may imply the presence of a

highly associative processing level which might generally be related to musical sensations and integrates widespread medial areas of the human cortex.

In sum, our results demonstrate temporally distinct stages in timbre perception which not only involve bilateral parts of the peri-auditory cortex but also medially situated regions of the human brain associated with emotional and auditory imagery functions.

4.2 Introduction

Investigators have made many important strides toward understanding the neural underpinnings of timbre perception (Goydke et al., 2004; Halpern et al., 2004; Menon et al., 2002; Pantev et al., 1998; Shahin et al., 2003). Timbre is a fundamental percept of sound and an indispensable key feature for the appreciation of sound quality other than pitch, duration, loudness, and spatial location (McAdams and Bigand, 1993). In particular, musical timbre corresponds to the quality of periodic and quasi-periodic tones that allows one to distinguish one sound from another when they are identical in pitch, intensity, location, and duration. Primarily, timbre corresponds to the spectral profile of a sound. It is considered the acoustic quality of auditory objects and is integral to the identity of acoustic sources (Griffiths and Warren, 2004). In terms of perception timbre is the tonal quality and texture which distinguishes aircraft noise from a siren, the voice of a friend from the voice of a stranger, and an instrumental tone from a pure tone. Thus, timbre represents psychologically relevant information which an animal or human being refers to while being exposed to any environmental setting. Even though the underlying acoustic correlates of pitch, loudness, and duration are well understood, timbre is a perceptual cue whose psychoacoustic underpinnings are not sufficiently stipulated.
Meanwhile, however, it has turned out that the recognition of timbre is not only associated with the processing of spectral but also temporal aspects of tonal stimuli (Balzano, 1986; Griffiths, 2001; Samson and Zatorre, 1994). While spectral flux and spectral centroid correspond to spectral levels of timbre, the attack time of a tone also determines its timbre (McAdams et al., 1995). Both variations at the spectral and the temporal level of a tone affect timbre. For example a tone played by a piano differs from a pure sine wave tone in that the piano tone is spectrally and temporally more complex. In particular, pure tones lack at least two important elements of timbre: the harmonics and the spectral flux. The latter is a variable which determines the magnitude of spectral envelope changes over the duration of a tone. Thus, timbre has to be considered a multidimensional property of sound (Toiviainen et al., 1998).

Thus, akin to laughter timbre is an acoustic domain which deserves particular scientific interest. First, researching timbre may prompt the revealing of interesting aspects of spectrotemporal integration during auditory processing. Furthermore, from

a neuroscientific perspective the examination of the neural underpinnings of timbre is also of obvious interest as the identification of the neural network mediating timbre information may allow researchers to learn more about the subtle functional architecture of the auditory system in the human brain, in particular pertaining to perceptual asymmetries as well as the interrelation between higher auditory functions and music. Timbre is an ideal issue to address in this context as it has turned out that untrained non-musicians are more amenable to perceive alterations in timbre unlike changes of pitch and duration. Interestingly, it has recently been borne out that differences in timbre patterns of tones also correspond to different emotional expressions (Goydke et al., 2004) which may explain why even hearing simple pieces of music elicits emotional sensations in listeners (Blood et al., 1999; Blood and Zatorre, 2001; Brown et al., 2004b).

Even though the investigation of the cerebral organization of music has been rapidly advancing during the last decade little is known so far about the precise network which mediates the complexity of music perception. Recent views propose that partly overlapping neural ensembles in both the right and the left hemisphere process the differential aspects of music (Altenmuller, 2001). However, there is some evidence which points to a stronger involvement of right hemisphere regions when nonmusicians listen to musical sounds (Tervaniemi et al., 1997). Akin to speech, music is not processed as monolithic entity but may be fractionated into functionally and neurally isolable components (timbre, pitch, rhythm, meter, and interval) (Besson and Schoen, 2001; Liegeois-Chauvel et al., 1998; Peretz and Coltheart, 2003). Accordingly, Seither-Preisler et al., (2003) reported a greater involvement of the right hemisphere in the processing of timbre tones characteristic of musical instruments and voiced portions of speech. This notion is also strengthened by clinical evidence which demonstrates that right hemisphere damage sustainedly affects the ability to properly process melodic contours, pitch and timbre information (Liegeois-Chauvel et al., 1998; Milner, 1962; Murayama et al., 2004) in particular after excision of the right temporal lobe (Samson and Zatorre, 1988; Zatorre, 1985). Another clinical observation observed deficient spectral processing in five patients after damage to the right temporoparietal region (Robin et al., 1990). In keeping with this finding other researchers noticed that damage involving the right temporal neocortex impairs the discrimination of spectral and temporal envelopes of tones characterizing musical timbre (Samson and Zatorre, 1994; Chobor and Brown, 1987). Furthermore, recent

imaging studies have shown that processing musical melodies preferentially drives right temporal sites (Overy et al., 2004; Samson and Zatorre, 1994). Buttressing evidence comes also from a study which investigated timbre perception in healthy volunteers by means of a dichotic listening paradigm and communicated a left ear advantage indicative of right hemisphere superiority (Brancucci and San, 1999). Collective results obtained from hemodynamic brain imaging show a more heterogeneous picture. In a PET-study (Platel et al., 1997) observed an involvement of right frontal sites during detection of timbral changes whilst two recent fMRI studies associated regions in both the left and right temporal lobe with the perception of timbre (Halpern et al., 2004; Menon et al., 2002).

However, lesion and brain imaging studies are less advantageous in that they do not provide a fine-grained temporal resolution which may uncover the subtlety of subsequent temporal processing steps associated even with simple sensory functions. Furthermore, one has to take into account that the precise relationship between hemodynamic responses collected by PET and fMRI on the one hand and neural activation on the other hand is still an insufficiently answered question (Marcar and Loenneker, 2004). To directly reveal the spatiotemporal pattern of brain activity during timbre perception it may be even more convenient to use the well established event-related brain potential (ERP) technique in combination with an electrical brain imaging tool meant to estimate ERP source location. For example, by means of ERPs it has recently been shown that the human cerebral auditory system automatically and rapidly categorizes inflowing instrumental tones based on subtle timbre differences reflected by an early negatively going waveform (Goydke et al., 2004).

Thus, auditory evoked brain potentials (AEPs) appear most suitable to investigate the spatiotemporal pattern of brain activity during timbre perception. As successfully evidenced the well established N1/P2 complex is most amenable to reflect small changes in auditory patterns. The N1/P2 can be elicited by the onset of almost any kind of sound, irrespective of its physical and or perceptual origin (Eggermont and Ponton, 2002; Langner et al., 1997; Naatanen and Picton, 1987, Vaughan and Ritter, 1970) with the N1 probably reflecting any detection of frequency changes in the auditory environment (Hyde, 1997; Menning et al., 2000). The treatment of the N1/P2 as unitary has recently been seriously questioned (Lutkenhoner and Steinstrater, 1998). This revised view has emerged as recent research on N1/P2 source

localization compellingly demonstrated that temporally and spatially distinct generators form the electrophysiological N1/P2 scalp potentials. For instance, the planum temporale (PT), a portion on the posterior supratemporal plane which accommodates auditory association cortex (Westbury et al., 1999), has been identified as the major source of the N100m (the magnetencephalographic equivalent of the N1) while neuromagnetic P200 appears to arise from the core region of the primary auditory cortex covering the medial part of Heschl's gyrus (Krumbholz et al., 2003; Lutkenhoner and Steinstrater, 1998; Pantev et al., 2001; Schulte et al., 2002). EEG evidence derived from point-dipole source modelling point to the existence of bilateral symmetrical sources of the N1 (and partly of the P2) situated in the secondary auditory fields of the supratemporal plane and the lateral temporal lobe (Bosnyak et al., 2004; Picton et al., 1999; Eggermont and Ponton, 2002; Scherg and von Cramon, 1986a; Scherg and von Cramon, 1986b).

However, the N1/P2 complex is not only capable of indicating the onset of a sound (Pratt et al., 2005). Apparently, the electrophysiological parameters of the N1/P2 complex (latency, amplitude, topography) are also responsive to subtle changes in the inflowing auditory stream. Accordingly, significant changes in ERP-components may reflect the greater synchrony of neural activity in the auditory cortex which echoes more complex auditory processing (Pantev et al., 1998). For instance, as a function of intensive auditory training (Bosnyak et al., 2004) reported a bilaterally enhanced P2 and an enlarged N1c in the right hemisphere where neural ensembles appear to be amenable to spectral information. One other ERP-study an enlarged N1/P2 complex recorded over temporal regions while subjects were engaged in pitch and timbre discrimination tasks (Jones et al., 1998). Due to the authors their results indicate the existence of a distinct timbre-sensitive mechanism which may be capable of detecting the onset of vocal information in a constantly modulated stream of sounds. More attesting evidence supporting the presumption of the right hemisphere as the major residence of neural networks preferring timbre and spectral information comes from two other EEG-studies (Auzou et al., 1995; Pardo et al., 1999) and one MEG-study (Seither-Preisler et al., 2003). However, the conclusions drawn from some of the above mentioned EEG-results which emphasize the role of right hemisphere sites remain elusive as they did not make the attempt to localize neural sources.

4.3 The present study

The novelty of the present study is the exploration of an alternative inverse solution method based on low-resolution electromagnetic tomography (LORETA) to estimate the putative neural sources of electrical brain activity in the context of musical processing (for details see Methods section). The LORETA approach has recently been successfully used in a number of EEG-studies on visual and auditory processing (Esslen et al., 2004; Gottselig et al., 2004; Laufer and Pratt, 2005; Mulert et al., 2002; Mulert et al., 2004b; Mulert et al., 2005; Pratt et al., 2005, Sinai and Pratt, 2003; Vitacco et al., 2002). Thus, one important scope of the present study is the exploration of spatio-temporal dynamics of timbre perception applying an established source estimation approach. As timbre is a multidimensional property of complex sound which differs from sine wave sound in spectral information (spectral flux, spectral centroid) and temporal information (attack time) we expect to find generators in bilateral temporal lobe structures. However, in keeping with the majority of aforementioned studies we predict a rightward asymmetry in neural activity as the right hemisphere has been widely acknowledged to play a pivotal role in musical processing (Janata et al., 2002; Tervaniemi et al., 1997). Notably, we consider it imperative to highlight that the LORETA approach to estimate source localization is not based on any a-priori assumptions determining the putative region of principal activation. Accordingly, by means of LORETA we might even be able uncover timbre-related activity in extra-auditory areas in which case we might gain new insights into higher auditory processes, particularly those associated with the processing of musical information. Furthermore, we predict to find differential spatial patterns for N1 and P2 component evoked by instrumental tones since recent EEG-studies also suggested that N1 and P2 may reflect the existence of distinguished tonal processing mechanisms residing in distinct brain sites (Bosnyak et al., 2004; Shahin et al., 2003).

However, we also regard it important to mention that the present study is not meant to distinguish between temporal and spectral processing steps during timbre perception. The present study has rather been undertaken to study event-related brain potentials evoked by instrumental tones (piano, trumpet, and violin) which incorporate a distinctive timbre compared to simple sine wave tones.

Finally, an additional aim sought to be addressed by the present study was the question to what extent the EEG-parameters are responsive to the cognitive processing mode. To explicitly test this issue all participants who partook in this study initially had to passively listen to the stimuli and subsequently to perform a discrimination task whilst hearing sine wave and complex instrumental tones. We hypothesize that N1/P2 parameters do not differ between active and passive listening conditions as the N1/P2 complex has been attributed to automatically flag any frequency change in the auditory environment while. On the other hand, we expect to find cognitively loaded late components associated with marked amplitude enhancements for the explicit discrimination tasks relative to the passive task which does not require any higher mental operation.

4.4 Materials and Methods

Subjects. We investigated 19 volunteers (25.7±0.9SE years of age, 7 females, 9 males). We excluded two participants due to strong muscle artefacts and another one due to insufficient performance in the task so that the remaining 16 volunteers entered data analysis. Participants had no formal musical training and they never played a musical instrument. According to the Annett-Handedness-Questionnaire (AHQ) (Annett, 1970) and HDT (Jancke, 1996) all subjects were consistently right–handed. No participant had any history of neurological or audiological disorders, nor did they report having absolute pitch. Subjects gave written consent in accordance with procedures approved by the local ethics committee and were paid for participation.

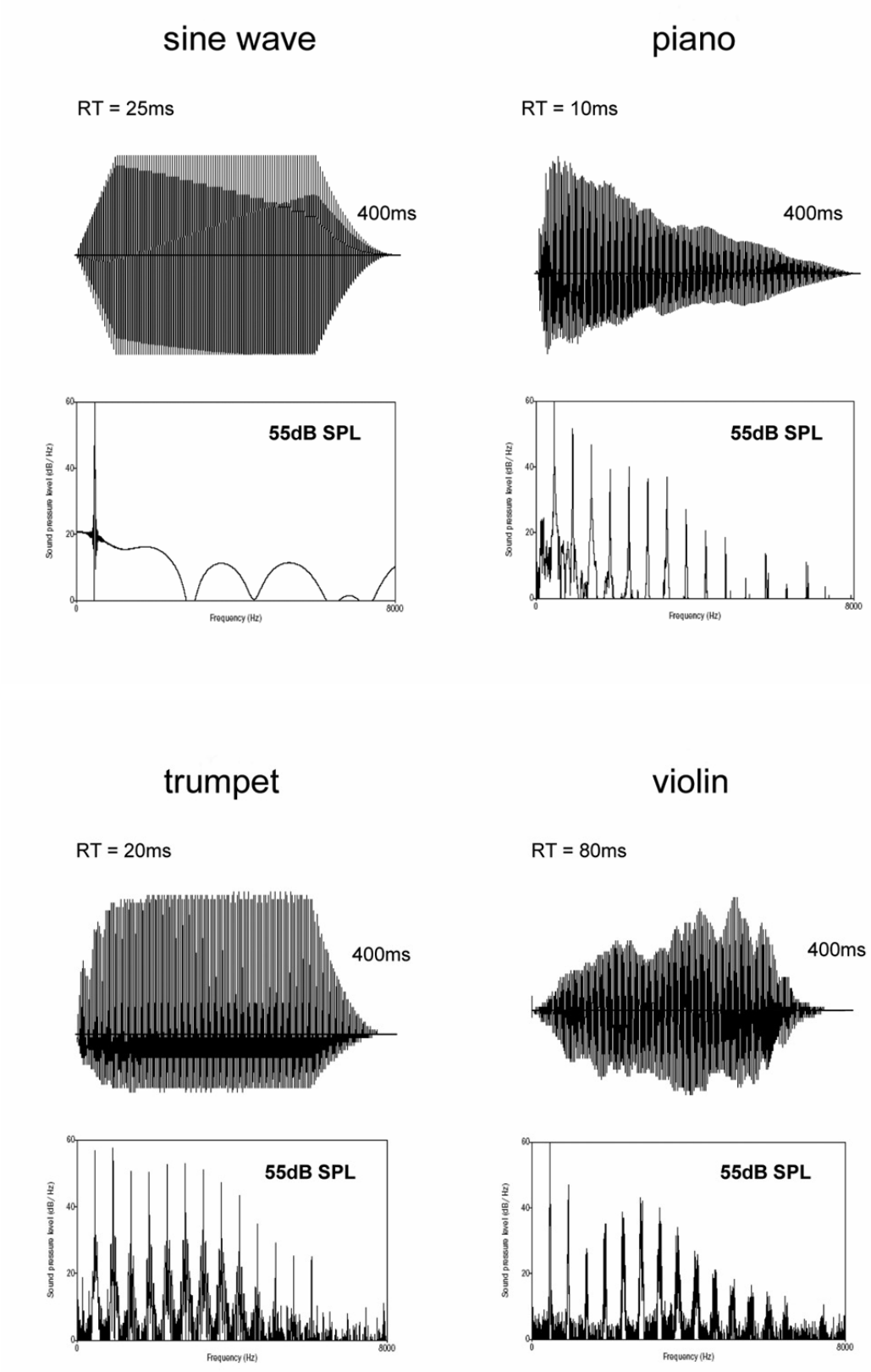


Fig.1. Temporal envelopes (upper panel) and spectra (lower panel) for the B4 sound stimuli.

Procedure. Subjects were placed in a comfortable chair at 110 cm distance from the monitor. We presented auditory stimuli binaurally using Hifi-headphones (55 dB sound pressure level). In total, the tonal stimuli comprised eight different computer-generated stimuli (MAGIX Audio Studio 03 deLuxe, Magix AG, Berlin, Germany). We presented subjects with sine wave tones, piano tones, trumpet tones, and violin tones at two different pitches (F4/349 Hz and B4/466 Hz, American notation)[6].

As apparent from Figure 1 tonal stimuli were equal in duration and sound pressure level, but differed in rise time and spectral composition. Instrumental tones slightly differed in spectral composition while sine wave tones simply consisted of a single frequency. The sounds all had a smooth onset and offset. The rise time to 60% of the maximal sound intensity was 10 ms for the piano tones, 20 ms for the trumpet tones, 25 ms for the sine wave tones and 80 ms for the violin tones.

All stimuli were matched for intensity by using Magix Audio Studio and re-checked by the PRAAT speech editing software[7]. In addition, we double checked sound pressure level by applying the *Voltcraft 329 digital sound level meter* (Conrad electronics, Wernberg, Germany).

Stimulation and recording of responses was controlled by the Presentation software (Neurobehavioral Systems, USA)[8]. Each session comprised four blocks. Prior to each block subjects were instructed to listen attentively to the stimuli and to perform task. Tasks were different for each block (1. passive listening, 2. pitch discrimination, 3. target detection (pure tone), and 4. target detection (one out of three instrumental tones)) and were balanced in that each session started with the 'passive listening' task.

Each of the four tone classes (sine wave, piano, trumpet, and violin) was presented 80 times in one experimental block and 320 times in the entire experimental session. Each tone class was made up of 40 F4 tones and 40 B4 tones. Thus, the total number of stimuli presented in the current study was 1280 (320 per block). Presentation of stimuli was matched in duration. Each tonal stimulus was presented for 400 ms followed by a 1600 ms Inter-Stimulus Interval (ISI).

---

[6] Audio files of experimental stimuli are available at http://www.psychologie.unizh.ch/neuropsy/home_mmeyer/ZEP

[7] http://www.praat.org

[8] http://nbs.neuro-bs.com/

Task. It has been shown that actively calling a subject's attention toward an auditory stimulus yields increased activation in auditory fields (Grady et al., 1997; Jancke et al., 2001; Zatorre and Binder, 2000). Therefore, we asked our participants to perform tasks of varying attentional demands and to direct their attention to particular acoustic aspects of the stimuli. For each participant each session started with a passive listening block. During this block participants were not required to perform a particular task but to press either the left or right button in alternating order after presentation of each single trial. During the second, third and fourth block participants engaged in differential discrimination task. By pressing one out of two buttons they had to indicate pitch, complexity, or instrument. We systematically balanced the order of explicit task blocks throughout the study.

EEG recording and analysis. The EEG (30 channels + 2 eye channels, subset of 10/10 system) was recorded with a sampling rate of 500 Hz and a band pass filter (0-100 Hz) using a Brainamplifier (Brainproducts, Munich, Germany). We applied sintered silver/silverchloride-electrodes (Ag/AgCl) and used the FCz position as reference. Electrode impedance was reduced to <10 k by Electrogel conductant. For all steps of digital EEG raw data processing we used Brain Vision Analyser software (Version 1.04, Brainproducts, Munich, Germany). The data was filtered at 1-45 Hz and eye movement artefacts (in some cases muscle artefacts) were removed using an independent component analysis (ICA) add on (Jung et al., 2000). The processed data was re-referenced to an average reference and epoched according to stimulus type (sine wave, piano, trumpet, or violin) and task (as described above) into 700 ms segments, including a 200 ms prestimulus baseline. Epochs were averaged for each subject according to stimulus type and task collapsing over F4 and B4 pitches and resulted in 16 grand averages. These grand averages were pooled into four stimulus data sets comprising all kind of tasks and four task data sets comprising all kind of stimuli. Thus, each of the remaining 8 grand averages contained 320 trials per subject.

In order to statistically confirm the difference between AEP components at Cz between sine

wave and instrumental tones we performed two tests. First, we ran two-tailed t-tests contrasting the voltage values for each sample point. The t-test against sine wave tones

were applied to the average of all instrumental tones as well as all instrumental tones separately. Secondly, AEP components were identified in the averaged data for each subject and for each of the 8 conditions by a peak detection algorithm. The N1 peak was determined as the absolute voltage minimum in a time window (TW) 50-150 ms after stimulus onset occurring at Cz. The P2 peak was determined as the absolute voltage maximum in a time window 150-250 ms after stimulus onset occurring at Cz. Peak data were subjected to a paired t-test to reveal amplitude differences between single conditions.

Source localization. We applied an inverse linear solution approach - LORETA (lowresolution electromagnetic tomography[9] to reveal the neural sources of event-related scalp potentials (Pascual-Marqui et al., 1994; Pascual-Marqui et al., 1999; Pascual-Marqui et al., 2002). This approach is advantageous in that - unlike conventional dipole fitting - it *does not* require assumptions about the number of dipoles. LORETA calculates the three dimensional distribution of electrically active neuronal generators in the brain as a current density value (∞μA/mm2) based on the recorded scalp electric potentials. LORETA provides a solution for the *inverse problem* by assuming that the smoothest of all possible activity distributions is the most plausible one to explain the data ("smoothness assumption"). Due to the smoothness constraint the most characteristic feature of the solution of the inverse problem is the low spatial resolution which conserves the location of maximal activity but with certain degree of dispersion (Mulert et al., 2004a). Here, we determined the current source density distribution for epochs of brain electrical activity on a dense grid of 2394 voxels at 7-mm spatial resolution. The localisation error of LORETA source identification may vary between 7 mm (Pascual-Marqui et al., 1994) and 11 mm (Phillips et al., 2002a; Phillips et al., 2002b). LORETA refers to a three-shell spherical model registered to the Talairach human brain atlas (Talairach and Tournoux, 1988). The source locations were therefore provided as (x,y,z) coordinates which in accord with standard conventions are relative to the intercommissural line (AC-PC line) in the horizontal (x), the anterior/posterior (y), and vertical (z) directions. The solution space was confined to the grey matter portion of the human cortex which rules out the option that subcortical tissue and white matter contribute to the solution. Numerous studies have been successfully using LORETA to delineate the

[9] http://www.unizh.ch/keyinst/NewLORETA/LORETA01.htm

neural sources of auditory evoked scalp potentials and presently LORETA can be considered an established way of providing reasonable estimations of neural generators (Mulert et al., 2005; Pascual-Marqui et al., 2002).

In order to verify the localisation of the N1 and the P2 component we calculated the LORETA current density value (∞μA/mm2) for sine wave tone and complex tones (average of the instrumental tones) AEP's within the 3D voxel-space. We used a transformation matrix with high regularisation (1e-3*(first eigenvalue)) to increase signal to noise ratio. The maxima of the current density distributions are displayed on a cortical surface model and transformed in stereotactic Talairach space (Talairach and Tournoux, 1988). In order to define adjusted time windows for the N1 and P2 component we calculated the global field power (GFP) derived from all EEG channels. The interval between the adjacent global field potential minima of the N1 and P2 components were taken as relevant time windows for illustration of current source density distribution and statistical contrasts of LORETA estimations. GFP based analysis yields the following time windows: sine wave tones, N1 96–172 ms, P2 172-270 ms; instrumental tones, N1 82-157 ms, P2 157-280 ms. We consider this approach an ideal observer-independent procedure to define relevant time windows for ERP and LORETA analysis[10]. Notably, the TWs of N1 and P2 for AEPs and LORETA estimations slightly differ as the AEP analysis only includes EEG signals sampled from Cz electrode while the LORETA estimation is based on signals collected from 30 electrodes.

We localized timbre specific activations considering the following current density distribution contrasts: complex tones vs. sine wave tones, piano tones vs. sine wave tones, trumpet tones vs. sine wave tones and violin tones vs. sine wave tones. The current density distribution for the complex tones was calculated from the averaged instrumental tone potentials. The contrasts are calculated from voxel by voxel t-tests. We assessed statistical significance corrected for multiple comparisons by means of a nonparametric randomization test (Nichols and Holmes, 2002).

Furthermore, we calculated a post-hoc region-of-interest (ROI) analysis to test whether activity significantly differing between pure and instrumental tones is larger in left or right hemisphere. We defined four ROIs in the auditory cortex (LH PAC, RH PAC, LH posterior SF, RH posterior SF) which comprised current source density (CSD) maxima. The landmarks of ROIs are determined by an automatic anatomical

[10] For a detailed description and further empirical validation of the 'GFP' approach visit 'www.psychologie.unizh.ch/neuropsy/home_mmeyer/ZEP'.

labelling procedure implemented in LORETA which is based on the position, size, and borders of Brodmann areas (BA 41/42 and BA 40) in stereotactic 3D Talairach space (Talairach and Tournoux, 1988).
Prior to SPSS (www.spss.com) based t-statistics the data was subjected to a one-sample Kolmogorov-Smirnov goodness-of-fit test which we used to decide if LORETA data comes from a population with a normal distribution. The results of this test (N1: LH posterior SF .645, RH posterior SF .601; P2: posterior SF .430, RH posterior SF .571) allowed us to generally subject data to paired samples t-tests.

4.5 Results

Behavioural data. The minimal requirements of task performance were set to 80% correct answers in each of the three explicit tasks. As mentioned above, only one person failed to fulfil this criterion and had to be excluded. Most of the participants performed more accurately than 95% correct answers in all tasks.

AEP waveforms time-locked to the onset of the tone. The analysis of event-related brain potentials yielded two major results. Firstly, we calculated difference waves between potentials collected during the passive listening block and the trials in the explicit task conditions. As apparent from Figure 2 the ERP responses to active and passive trials do not differ significantly in the typical N1/P2 time window. Based on this finding we collapsed ERP's obtained from passive and active tasks for all further steps of analysis to magnify statistical power. However, at CZ all active tasks elicited a significant positive waveform which differed considerably from the passive task in the typical P3 time window. This difference between active and passive tasks was verified by paired t-tests. The difference between the passive task and the complexity task was significant between 260 and 424 ms ($p<0.01$), the passive task and the instrument task differed significantly between 272 and 444 ms ($p<0.01$), and the difference between the passive task and the pitch task was significant between 210 and 424 ms ($p<0.01$).

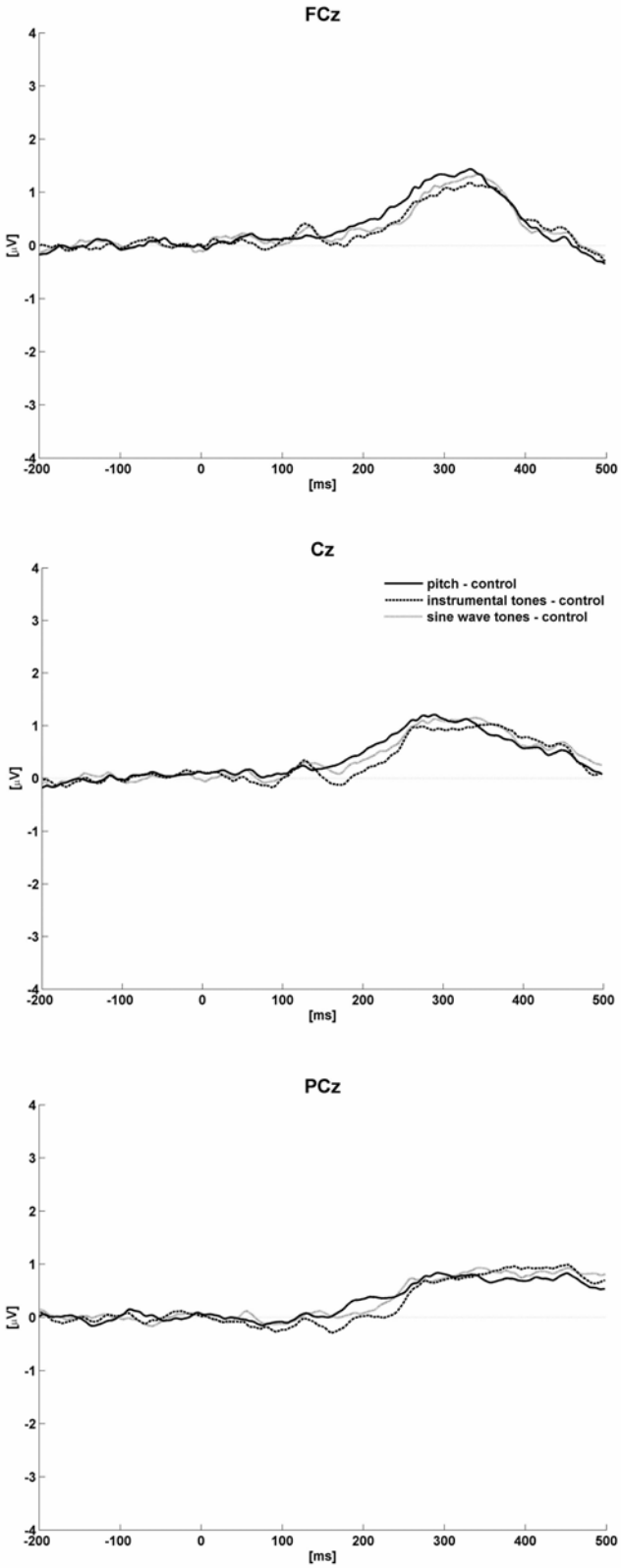


Fig. 2. Grand average ERP difference waveform for **sine wave task, instrumental task, and pitch task vs. control** task obtained from three scalp electrodes (FCz, Cz, PCz).

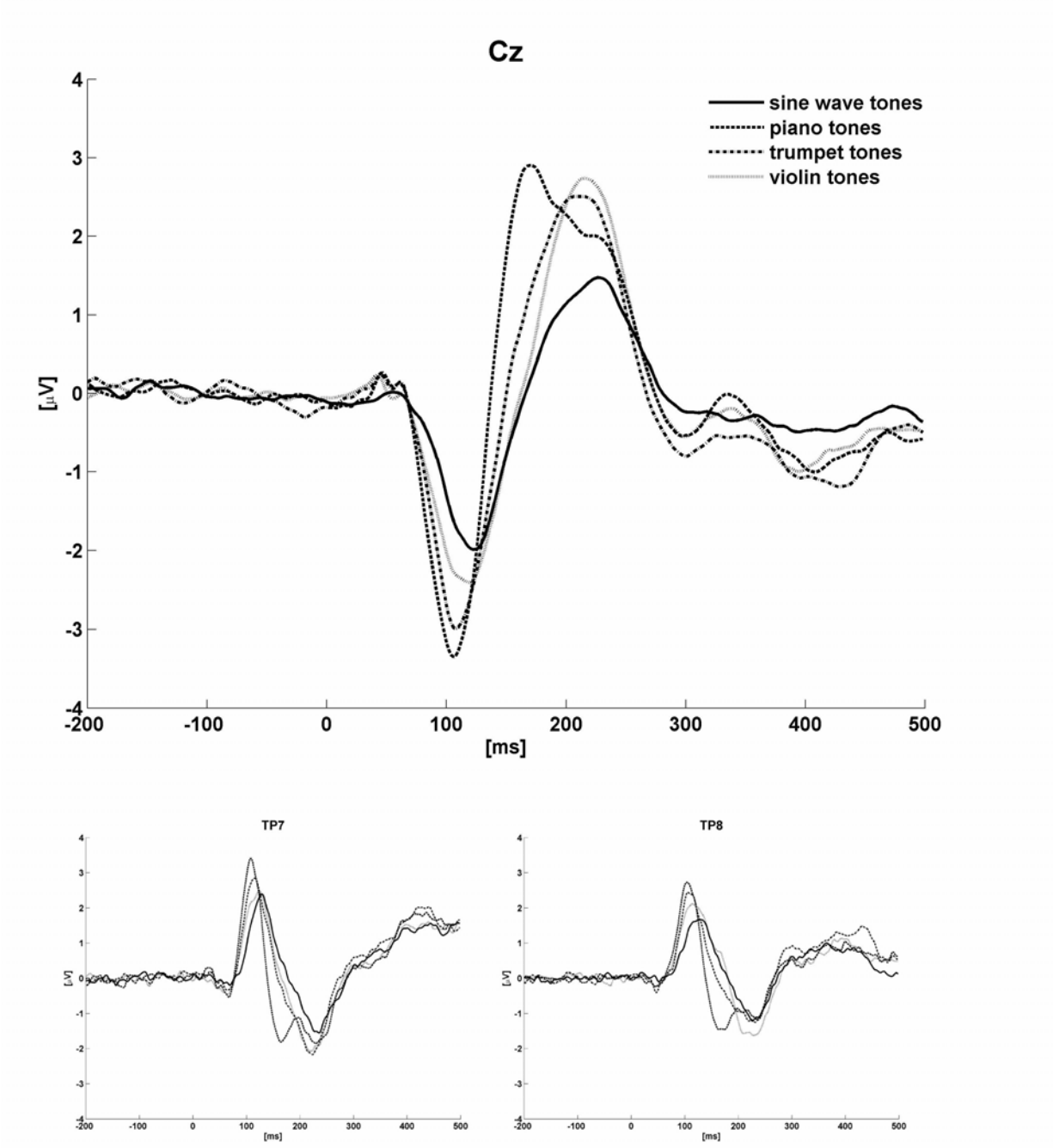


Fig. 3. Grand average ERP difference waveform for **sine wave task, instrumental task, and pitch task vs. control** task obtained from three scalp electrodes (FCz, Cz, PCz).

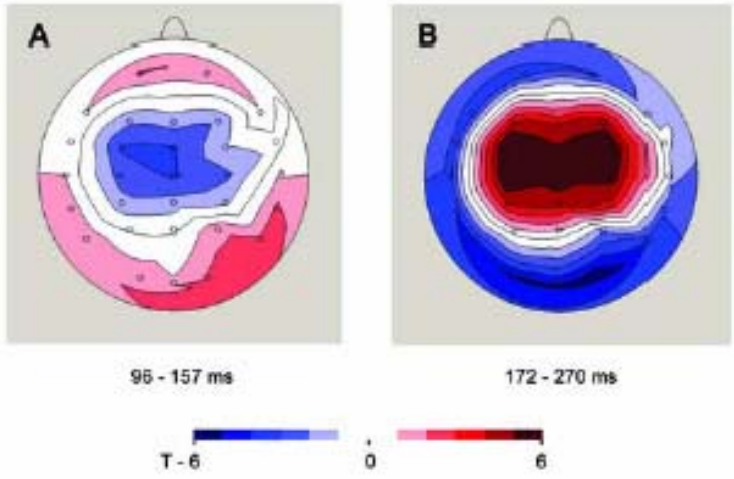


Fig. 4. Instrumental vs. sine wave tones: topographic scalp maps of t-values collected during the N1 (4A) and P2 (4B). Blue colours denote negative t-values and red colours depict positive t-values.

Secondly, as depicted by Figure 3, at CZ electrode a paired t-test uncovered a clear amplitude difference ($p<0.01$) between sine wave tones and instrumental tones in N1 (TW: 78-116 ms) and P2 (TW: 146-228 ms) components[11]. Results of t-tests on peak data derived from CZ for the N1 and P2 amplitude electrode revealed significant differences between sine wave tones vs. piano tones (N1, $p<.000$; P2, $p<.000$), sine wave tones vs. trumpet tones (N1, $p<.000$; P2, $p<.000$), and sine wave tones vs. violin tones (N1, $p<.001$; P2, $p<.000$).

Source localization. Table 1 presents coordinates as well as brain regions associated with the current density maxima separately for sine wave tones and instrumental tones obtained from the GFP based N1 and P2 time window.

For the N1 time window current source density maps of instrumental and sine wave tones indicate that left and right posterior Sylvian areas, namely the primary auditory cortex (PAC) and the planum temporale (PT) as well as the right deep frontal operculum encroaching onto the anterior insula contribute to timbre and pure tone perception. With regard to the P2, LORETA uncovered enlarged current source density distributions for complex relative to simple tones. Mainly, the left and right

[11] Figure 4 shows that the N1 and P2 differences are prominent not only at Cz but also at adjacent electrodes.

auditory cortex and medial frontal gyrus including the cingulate and the paracentral lobe constitute the P2 component irrespective of tone complexity.

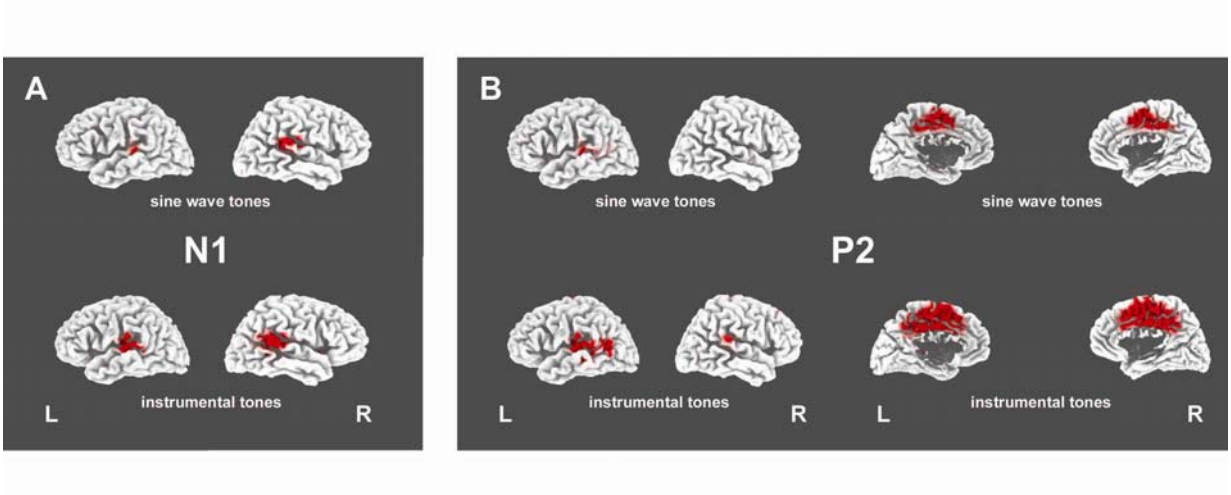


Fig. 5. Grand average (n=16) three dimensional LORETA-based current density maxima for AEP components: A, N1 component (time window: sine wave 96-172 ms; instrumental 82-157 ms; threshold, 0,001 prop. A/m2). B, P2 component (time window: sine wave 172-270 ms; instrumental, 157-280ms; threshold, 0,0008 prop. A/m2).

Table I: Current density maximum [μA/mm$_2$]*10^{-3} activations in response to sine wave and instrumental tones.

Component	Condition	Brain region	Current density value	Left Hemisphere			Right Hemisphere		
				x	y	z	x	y	z
N1	**Timbre**	STG BA 42	1.428		-		60	-32	15
		STG BA 41	1.277	-59	-25	8		-	
		INS BA 13	1.042		-		46	10	13
P2		Cingulate BA 24/6	1.272	-3	-11	50		-	
		STG BA 42	1.083	-59	-32	8		-	
		STG BA 22	0.824		-		60	-39	15
N1	**Sine wave**	STG BA 42	1.121		-		60	-32	15
		STG BA 41	1.024	-59	-25	8		-	
		INS BA 13	0.923		-		46	10	1
P2		Cingulate BA 24/6	1.123		-		4	-4	43
		STG BA 42	0.81	-59	-32	8		-	
		STG BA 22	0.724		-		60	-39	15

LORETA t-statistics. The statistical contrasts yield consistently increased activity for complex instrumental relative to pure tones during both the N1 and P2 TW.

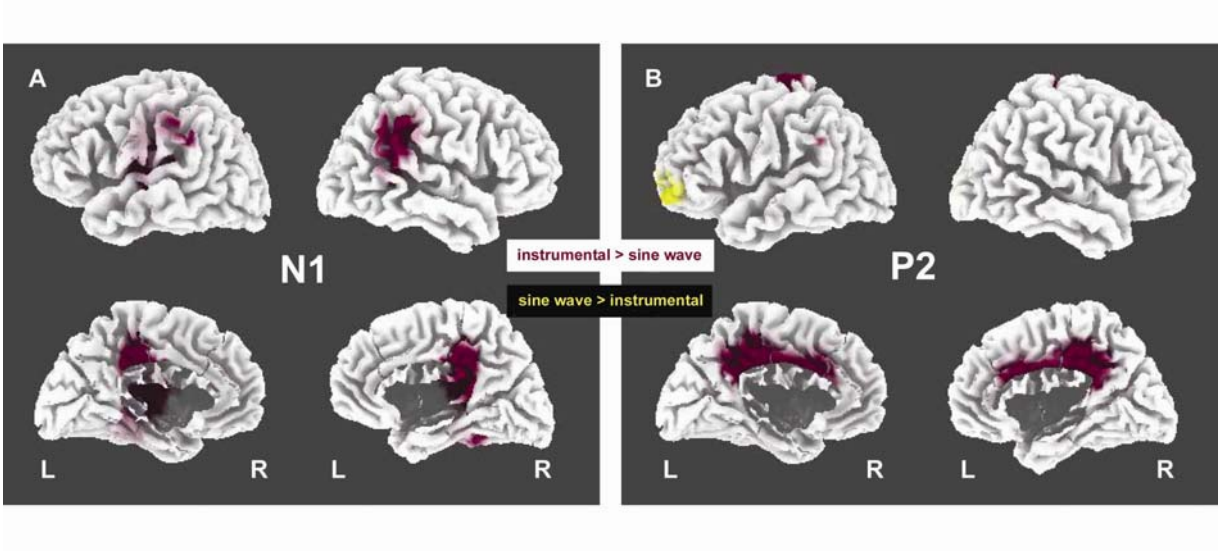


Fig. 6. Grand average (n=16) three dimensional LORETA-based statistical nonparametric maps for **sine wave versus instrumental** tones contrast: A, N1 component (time window: sine wave 96-172 ms; **instrumental** 82-157 ms; threshold, t-value 4.3). B, P2 component (time window: sine wave 172-270 ms; **instrumental**, 157-280ms; threshold, t-value 4.3).

For the N1 TW the t-test reveals clear activity in right the posterior SF covering auditory association cortex even encroaching onto the inferior parietal lobe (IPL), in particular the supramarginal gyrus (SMG). In the left hemisphere LORETA highlighted less significant differences between complex and pure tones in the primary auditory cortex and posterior portions of the peri-sylvian region lining the IPL. Additionally, N1 generators more extensively seem to originate the posterior cingulate cortex. For the P2 TW analysis of timbre more strongly recruits left posterior SF (SMG), anterior and posterior cingulate cortex also encroaching onto the paracentral lobe. Notably, an area of the inferior frontal gyrus more strongly supported the perception of simple compared to complex tones. The post-hoc ROI analysis confirms stronger right over left posterior SF (BA 40) engagement in timbre perception for the N1 ($p<0.01$) and clearer left over right superiority in the P2 TW for the same brain region ($p<0.05$). For the primary auditory cortex (BA 41/42) we did not find a significant difference between left and right hemispheric neural activity in the N1 TW ($p=.17$). Evidently, this result implies a time dependency of hemispheric

dominance during timbre perception. While we uncovered a prominent rightward lateralization in the posterior

Table II: LORETA t-statistics for maximum activations obtained from contrasts between sine wave and instrumental tones. Table 2 displays t-values for current density maxima, threshold for significance ($p<0.05$, corrected for multiple comparisons) is at a t-value of 3.95.

Component	Contrast	Brain region	T-value	Left Hemisphere			Right Hemisphere		
				x	y	z	x	y	z
N1	Timbre > Sine wave	post. SF BA 22/40	8.23				53	-39	29
		PCC BA 31	7.9	-3	-25	36		-	
		post. SF BA 22/40	6.85	-59	-39	36		-	
		STG BA 41	6.65	-45	-25	8		-	
P2	Timbre > Sine wave	PCC BA 23/31	6.66		-		4	-39	29
		ACC BA 24/33	6.45		-		4	10	29
		post. SF BA 22/40	4.51	-59	-44	29			
P2	Sine wave > Timbre	IFG BA 10	-4.93	-38	60	-6			

SF at an early processing stage (N1 TW), we observed a significant shift of neural activity during the subsequent processing stage (P2 TW) to the contralateral site. Based on this distinct spatiotemporal pattern it is plausible to reason that N1 and P2 may reflect different stages of timbre processing. We will more thoroughly address this issue in the Discussion. Furthermore, we ran a comparison of single instrumental tones (piano, violin, and trumpet) and sine wave tones which revealed that the LORETA pattern did not considerably differ between differential timbres[12].

[12] For an illustration of separate contrasts (piano, violin, and trumpet vs. sine wave tones) visit 'www.psychologie.unizh.ch/neuropsy/home_mmeyer/ZEP'.

4.6 Discussion

We recorded EEG in healthy participants with normal hearing skills to examine the temporal aspects of timbre perception. In combination with a distributed source model - (LORETA) – an electrical brain imaging method to estimate the neural sources of EEG signal – we revealed different succeeding temporal stages of timbre processing which correspond to distinct neural sources in auditory and non-auditory regions.

Event-related brain potentials

In the present study, the major parameters of auditory evoked potentials (N1/P2) turned out to be insensitive to task demands. All participants were asked to either passively listen to the tones or explicitly discriminate acoustic features of stimuli. Data analysis did not reveal any noticeable differences in latency, polarity, or amplitude between AEP components for active vs. passive participation. However, as revealed by the difference wave between active vs. passive tasks we observed a marked positivity peaking around 300 ms for all trials which are associated with explicit discrimination. Thus, we may conclude that early stages of timbre perception occurred automatically while the P3-like component reflects task-relevant neural effort associated with any, probably unspecific, cognitive processes. Exactly the same pattern of responses has recently been reported by Pratt et al. (2005) which lends confidence to our observations. Typically, the P300 component is easily elicited by rare events embedded in an oddball paradigm which requires the participant to categorize or discriminate auditory of visual events (Mulert et al., 2004b). The present study, however, did not utilize an oddball paradigm so that we are not able to legitimately label the late positivity a P300. We rather prefer to use the notion ‘P3-like component’ or ‘P3b’ which is highly likely to mark task-related categorization processes and stimulus evaluation (Friederici et al., 2001; Goydke et al., 2004).

As a function of coustic complexity we observed significantly increased N1/P2 amplitudes for all types of instrumental tones as compared to sine wave tones. This finding fully concurs with our predictions and partly replicates a previous ERP study which reported enhanced electric P2 amplitude for piano and violin tones relative to pure tones (Shahin et al., 2003). In another EEG-study Shahin and colleagues also observed slightly larger P2 amplitude for violin and piano tones across children with and without musical training (Shahin et al., 2004). The novelty of our study is that we additionally observed significantly magnified N1 amplitude for instrumental tones irrespective of the particular instrument. It is most perspicuous to assume that spectro-temporal complexity of instrumental tones account for the observed difference as N1 and P2 amplitude have been related to indicate any discrete change in any dimension of the auditory environment (Eggermont and Ponton, 2002; Vaughan and Ritter, 1970).

The finding of differences in latency of N1 and P2 peaks for instrumental relative to pure tones also deserves discussion. We observed shortest latency times for piano and trumpet tones and relatively longer latency times for violin and pure tones. Notably, as apparent from Figure 1 piano and trumpet tones have shorter rise times, that is duration from physical onset of acoustic signal to 60 % of maximum. This is because instrumental tones do not have a consistent plateau. Notably, differences in attack ramps do apparently not account for ifferences in the N1 and P2 amplitude as artificially generated sine wave tones show lowest mplitude but mean rise time relative to piano, violin, and trumpet tones.

Low resolution source estimation of auditory evoked potentials

The major target of the present study was to reveal the neural networks underlying the auditory evoked responses to simple and timbre tones. For this purpose we used a source estimation approach (LORETA) which is not based on any *a priori* assumptions on number and locations of generators and is hence supposed to uncover even widely distributed networks which shape scalp recorded potentials (Pascual-Marqui et al., 1994; Pascual-Marqui et al., 1999). As suggested by the

LORETA current source density results the N1 and P2 component originate from partly overlapping regions which correspond to auditory cortex.

Current source density (CSD) maps. As the conventional analysis of AEPs evinced clearly
distinguishable N1 and P2 components we performed distinct location of current density
maxima for N1 and P2. Irrespective of tonal complexity the current source density distribution during the N1 time window indicates neural origins in mid and posterior temporal sites with the right hemisphere playing a considerably stronger role. In analogy to ERP amplitude superiority observed for timbre relative to sine wave tones the source estimation for timbre tones uncovers a magnified distribution of neural excitement in left posterior Sylvian region. Basically this finding is in keeping with results delivered by conventional point-dipole analysis of the neuromagnetic auditory evoked field which suggest that peak N1m arises from the PT (Lutkenhoner and Steinstrater, 1998). Apparently, N1 generators partly seem to reside in the vicinity of the deep frontal operculum and/or the anterior insula. Even though these regions are cytoarchitectonically distinct (Bamiou et al., 2003) we cannot precisely say as to whether the signal originates from the first region, from the latter region or conjointly from the two regions. Hemodynamic responses from this brain region in the context of tone and music perception have recently been reported by neuroimaging studies (Brown et al., 2005; Brown et al., 2004a; Wong et al., 2002; Zatorre et al., 1994). At least one study which also applied the LORETA solution reported increased current source density in the right insula in the context of an auditory discrimination paradigm (Gottselig et al., 2004). Brown and colleagues even propose that this region plays a key role in the human song system (Brown et al., 2004a). Based on our data we cannot associate any specific functional role with the right frontal operculum but we take this converging data as compelling evidence that validates the LORETA estimation. With regard to the P2 component the CSD maps suggest three principal brain regions as plausible sources. Firstly, excited areas cover the left and right posterior Sylvian fissure in the vicinity of the planum temporale and the parietal operculum. Interestingly, in contrast to the N1 source distribution we noted a preponderance of activity in the left temporal lobe. Secondly the current densities arise from the medial part of the fronto-parietal zone (covering the cingulate gyrus,

the paracentral lobe and the medial frontal gyrus). Generally, the spatial extent of neural excitement is enhanced for the complex compared to the pure tones which again can be interpreted as reflecting the increased acoustic complexity of the instrumental tones. The finding of estimated posterior temporal and parieto-opercular activity underlying both the N1 and P2 lends credibility to our source localization as it is established knowledge that these peri-sylvian regions subserve auditory functions (Adriani et al., 2003; Griffiths and Warren, 2002; Jancke et al., 1994; Jancke et al., 2002; Jancke and Steinmetz, 1993; Meyer et al., 2005; Specht and Reul, 2003; Zaehle et al., 2004).

Effect of instrumental tones. Our LORETA analysis yields consistently increased activity for complex instrumental relative to pure tones during both the N1 and P2 time window in several distinct regions which we will discuss in turn.

Firstly, our distributed source model confirms stronger involvement of bilateral auditory areas in timbre perception which generally converges with recent brain imaging studies (Halpern et al., 2004; Menon et al., 2002).

Principally instrumental tones differ from pure sine wave tones in that the former incorporate more complex spectral and temporal information. Thus, it comes as no surprise that our study revealed more substantial bilateral engagement of primary and secondary auditory regions for instrumental tones compared to sine wave tones (Griffiths, 2001). This finding holds particularly for the N1 component which has been primarily associated with auditory processing (Eggermont and Ponton, 2002). However, as mentioned above this study has not been undertaken to tease apart temporal and spectral cues as timbre is considered a multidimensional feature of sounds.

Even though it is still unsettled to what extent the left and right auditory cortex may functionally differ, recent evidence strengthens the view that left auditory fields are preferentially driven by rapid temporal processing (Zaehle et al., 2004) while right auditory areas are more sensitive to spectral processing (Bosnyak et al., 2004; Zatorre and Penhune, 2001). However, as recently argued by an alternative model there are no absolute differences, merely relative differences in hemispheric specialization (Hickok and Poeppel, 2004). The principle of preferential processing of spectral or temporal features at the level of auditory cortex reflects relative computational specialisation after initial signal processing rather than fundamental

differences in signal processing between the two hemispheres. Basically our data concur with these two views in that both the left and right primary auditory and adjacent associative cortex engage in timbre perception, with the right hemisphere apparently playing a more important role. According to clinical data the right temporal lobe is particularly essential for analysis of slow temporal and spectral information involved in musical timbre (Samson and Zatorre, 1994) which indicates that timbre processing mainly depends on the integrity of the right temporal cortex. Thus, our data are in principal accord with the notion that both spectral and temporal features of sound are extracted from the acoustic signal in the primary auditory cortex and integrated in the secondary and tertiary areas to form mental representations of auditory objects and their spatial locations. We add evidence to the view that timbre perception cannot be considered an exclusive right temporal lobe function as bilateral recruitment of primary and secondary auditory cortex has also been shown in recent brain imaging studies (Halpern et al., 2004; Menon et al., 2002). Even though our LORETA results imply a high degree of accordance with functional imaging data we are fully aware that the spatial resolution provided by LORETA is the factor 2.5 – 3 worse relative to present fMRI scanning protocols. Hence our interpretation of the data is limited by the constraint that LORETA is an estimation of neural sources generating the EEG signal. Thus, we think the LORETA results do not justify a detailed discussion on the possible specific functions the distinct peaks may reflect but allows at least the discussion of asymmetric hemispheric involvement in timbre perception. Therefore, the finding of significantly stronger recruitment of the right posterior Sylvian region and the IPL requires an elaborate commentary. Generally, this finding converges with former clinical, behavioural, and electrophysiological studies which implied a particular sensitivity of the human right hemisphere for timbre and spectral processing (Auzou et al., 1995; Brancucci and San, 1999; Milner, 1962; Pardo et al., 1999; Samson and Zatorre, 1994) but enhances the present understanding in that it strongly supports the present knowledge about the functional role of the PT as this region has been associated with higher order processing of auditory input. Recent findings give argument in favour of the left PT mediating the analysis of rapidly changing acoustic spectra available in speech and nonspeech stimuli (Jancke et al., 2002; Meyer et al., in press; Zaehle et al., 2004). Complementarily, a PET-study by Belin and colleagues demonstrated a functional rightward PT asymmetry to slow relative to rapid formant transitions (Belin et al.,

1998). Corroborating evidence also stems from fMRI-studies that implicate a functional designation of computing slow prosodic modulations to the right PT (Meyer et al., 2002; Meyer et al., 2004).

In addition to PT activity LORETA also exposed a considerably stronger difference between simple and complex tones in the posterior SF including the SMG/PP. Based on the spatial resolution provided by LORETA we are not capable of dissociating to what extent the enhanced responses emerge in SMG or PP, but we propose to consider the right posterior SF (PP/SMG) as essentially mediating acoustic information associated with musical timbre. Lesion and imaging studies conjointly indicate that the SMG is also attributed to auditory processing, in particular short-term memory for processing of pitch and musical information (Celsis et al., 1999; Clarke et al., 2002; Gaab et al., 2003; Gaab and Schlaug, 2003). However, the majority of these studies report stronger government of left SMG while we discovered a functional rightward asymmetry. Interestingly, this finding receives indirect support by *in vivo* MR morphometry which has successfully demonstrated a structural rightward asymmetry of the planum parietale (PP) in right-handed man (Jancke and Steinmetz, 1993). Anatomically, PP is contiguous with the posterior temporal plane, overarched by the SMG, deeply buried in the IPL and part of the posterior ascending branch of the SF. As the PP partly overlaps with cytoarchitectonic area Tpt which has been classified as auditory parakoniocortex that extends contiguously from the STP into parietotemporal operculum (PTO) (Galaburda and Sanides, 1980) its involvement in auditory processing is perspicuous (Adriani et al., 2003; Maeder et al., 2001; Meyer et al., 2005; Warren et al., 2005). Admittedly, our data cannot be straightforwardly explained by the present knowledge on the functional role of this particular region as it has been demonstrated that the bilateral PTO preferentially responds to changes of sound source motion (Warren et al., 2002; Zatorre et al., 2002). However, recent functional imaging studies also noted an extensive engagement of the bilateral PTO while participants heard stationary sounds which could not have been spontaneously associated with meaning (Lewis et al., 2004; Meyer et al., 2005). Remarkably, Zatorre and colleagues provided a proposition derived from a PET-experiment which may reconcile the discrepant findings (Zatorre et al., 1999). Due to their findings it appears that modulations of PTO activity are very familiar when subjects attended to spatial **or** spectral features of the auditory input. Thus, it may be plausible to assume that the bilateral PTO is not specifically devoted to any particular function but may

serve any complex auditory processing, that is any temporal, spectral, and spatial pattern available in acoustic cues (Griffiths et al., 2004). Actually, Griffiths proposes a view which defines an acoustic experience as an auditory object as long as it requires the integration of temporal, spectral, and spatial patterns (Griffiths and Warren, 2004). In particular, timbre can be considered the most important attribute which might characterize the perceptual quality of an auditory object. In Griffith's view auditory objects – like timbre - are not analysed in either the spatial, temporal or spectral domain, but should be regarded as higher order patterns which recruit distributed parieto-temporal networks. Hence, the notion of the PTO as mainly governing sound source localization does not invalidate our results as the perception of instrumental tones imperatively requires brain regions, namely the PTO - devoted to the analysis of auditory objects, which is the analysis of temporal, spectral, and spatial sound qualities. Basically, the particular involvement of auditory association areas during complex tones also concurs with newly introduced delineations of the microstructure of auditory cortex. Based on cytoarchitectonical observations, recent neuroanatomical work in primates suggests a hierarchical organization of the auditory system with separate primary auditory 'core' areas and encompassing secondary 'belt' and tertiary 'parabelt' areas (Kaas et al., 1999; Kaas and Hackett, 1998; Kaas and Hackett, 2000). This structural division finds support from functional studies in monkeys (Rauschecker et al., 1995; Rauschecker et al., 1997; Rauschecker, 1998) and man which propose a functional correspondence with the core areas processing simple tones and the belt and parabelt areas being preferentially driven by complex sounds, nonspeech vocalizations, and even speech (Scott and Johnsrude, 2003; Specht and Reul, 2003; Wessinger et al., 2001). Here, the LORETA source estimations for the N1 component during timbre perception in PTO situated posterior to the primary auditory cortex are in principal keeping with the hierarchical organization of mammalian auditory areas. To sum up, our main result of a bilateral but right preponderant PTO government for timbre analysis finds ample support in the literature. Based on an extensive review of lesion and imaging studies Samson concludes that timbre perception is predominantly accommodated by right temporal areas with the role of the left hemisphere still remaining controversial (Samson, 2003).

Our LORETA estimation also uncovered a recruitment of extra-auditory areas during timbre perception. In particular, LORETA identifies larger N1 sources for instrumental tones in the PCC and extensively distributed P2 sources along the entire medial swathe with local maxima in the anterior and posterior cingulate. In the context of electrical brain imaging this result is by all means novel and deserves magnified interest. Perception of instrumental tones apparently recruits brain regions which are not tied to sensory processing. Principally, this is a plausible finding as it is not likely to assume that auditory cortex works in isolation but rather interacts with other areas concerned with action and cognition (Griffiths and Warren, 2004).

Electromagnetic evidence links the activation of cingulate activity to the unspecific analysis of auditory-object-feature detection (Knosche et al., 2005). Noteworthy, the activity spreads along the cingulate as a function of time. While we observed confined activity in the vicinity of PCC during the N1 time window, the ACC and PCC are conjointly responding during the P2 time window. The functional role of PCC is hitherto unspecified even though there is some evidence that awards the PCC an essential role in the modulation of selective attention (Damasio et al., 2000). In addition, the PCC has been coined as the “mind’s eye” since this area has been consistently shown to take part in visual imagery sets or situations (Fletcher et al., 1995). In the context of the present study, we conjecture that listening to instrumental tones may automatically evoke visualized images while hearing sine wave tones does not. The ACC, however, has so far been associated with emotional imagery (Bush et al., 2000; Phan et al., 2002), emotional monitoring, and attentional functions. Thus, it cannot be ruled out that unspecific emotional associations tied to the instrumental tones may account for the ACC activity in the present study. At least, this interpretation is in line with results of an electromagnetic source localization study which reports ACC engagement during music perception (Knosche et al., 2005). Based on this observation we may conjecture that the N1 reflects sensory-based auditory analysis whereas P2 is more densely related to unspecific cognitive and emotional processes.

Imperatively, based on our source localization we agree with Lutkenhoner and Steinstrater (1998) that the treatment of the complex N1-P2 as unitary, which was common practice in early electrical studies (Hyde, 1997), is evidently a questionable procedure.

In any case we conclude that the cortical structures we have been discussing could be considered a distributed network consisting of auditory and extra-auditory areas which are indispensable for the operation of timbre.

Once data collected from brain imaging are discussed in comparison to data obtained from electrophysiological and electromagnetic studies it should be born in mind that there are still open questions pertaining to the equivalent interpretation of hemodynamic and neural responses. Even though spatial resolution of electrophysiologically-based localization methods is inferior to modern neuroimaging techniques we consider the LORETA estimation a reasonable approach to investigate the neural organization of auditory perception for two principal reasons. First it allows to more precision in elucidating the temporal aspects of distributed network activations. Secondly, it permits direct measurement and analysis of neural activation. This is advantageous as hemodynamically-based methods are not capable of directly recording neural activity and the interpretation of BOLD signal changes are hampered by our insufficient present understanding of the neurovascular relationship to neural activity (Marcar and Loenneker, 2004).

4.7 Conclusion

The current study provides novel insight on the neural organization of timbre perception by describing distinct temporal processing stages based on EEG-data. Furthermore, by applying an innovative EEG source estimation which is not based on any *a priori* assumptions on number and locations of generators our results point to the existence of distinctly distributed networks which may underlie the N1 and P2 components brought on by instrumental timbre. First, the results for the N1 time window complement a body of data indicating a privileged role of right hemisphere mechanisms in music perception. More specifically our study extends these indications by revealing an involvement of primary auditory and right secondary auditory regions as principally governing the analysis of complex auditory objects. Additionally, the P2 source estimation suggests that extra-auditory areas which may be more strongly attributed to unspecific cognitive and emotional functions are also

part of the neural network binding together sensory and higher order associative areas to cooperate in basic music perception.

Acknowledgements

The authors wish to thank Roberto Pascual-Marqui and Adam McNamara for helpful comments on the manuscript. During the preparation of this manuscript Martin Meyer and Simon Baumann were supported by Schweizer National Fonds (Swiss National Foundation) SNF 46234101, 'Short-term and long-term plasticity in the auditory system'.

4.8 References

4.9

Adriani, M., Bellmann, A., Meuli, R., Fornari, E., Frischknecht, R., Bindschaedler, C., Rivier, F., Thiran, J.P., Maeder, P., Clarke, S.,2003. Unilateral hemispheric lesions disrupt parallel processing within the contralateral intact hemisphere: an auditory fMRI study. Neuroimage. 20 Suppl 1, S66-S74

Altenmuller, E.O.,2001. How many music centers are in the brain? Ann. N. Y. Acad. Sci. 930, 273-280

Annett, M.,1970. A classification of hand preference by association analysis. Br. J. Psychol. 61, 303-321

Auzou, P., Eustache, F., Etevenon, P., Platel, H., Rioux, P., Lambert, J., Lechevalier, B., Zarifian, E., Baron, J.C.,1995. Topographic EEG activations during timbre and pitch discrimination tasks using musical sounds. Neuropsychologia 33, 25-37

Balzano, G.J.,1986. What Are Musical Pitch and Timbre. Music Perception 3, 297-314

Bamiou, D.E., Musiek, F.E., Luxon, L.M.,2003. The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res. Brain Res. Rev. 42, 143-154

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M.C., Samson, Y.,1998. Lateralization of speech and auditory temporal processing. J. Cogn Neurosci. 10, 536-540

Besson, M. and Schoen, D.,2001. Comparison between language and music. Ann. N. Y. Acad. Sci. 930, 232-258

Blood, A.J. and Zatorre, R.J.,2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc. Natl. Acad. Sci. U. S. A 98, 11818-11823

Blood, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C.,1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. Nat. Neurosci. 2, 382-387

Bosnyak, D.J., Eaton, R.A., Roberts, L.E.,2004. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. Cereb. Cortex 14, 1088-1099

Brancucci, A. and San, M.P.,1999. Laterality in the perception of temporal cues of musical timbre. Neuropsychologia 37, 1445-1451

Brown, S., Ingham, R.J., Ingham, J.C., Laird, A.R., Fox, P.T.,2005. Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. Hum. Brain Mapp. 25, 105-117

Brown, S., Martinez, M.J., Hodges, D.A., Fox, P.T., Parsons, L.M.,2004a. The song system of the human brain. Brain Res. Cogn Brain Res. 20, 363-375

Brown, S., Martinez, M.J., Parsons, L.M.,2004b. Passive music listening spontaneously engages limbic and paralimbic systems. Neuroreport 15, 2033-2037

Bush, G., Luu, P., Posner, M.I.,2000. Cognitive and emotional influences in anterior cingulated cortex. Trends Cogn Sci. 4, 215-222

Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J.P., Berry, I., Nespoulous, J.L., Chollet, F.,1999. Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. Neuroimage. 9, 135-144

Chobor, K.L. and Brown, J.W.,1987. Phoneme and timbre monitoring in left and right cerebrovascular accident patients. Brain Lang 30, 278-284

Clarke, S., Bellmann, T.A., Maeder, P., Adriani, M., Vernet, O., Regli, L., Cuisenaire, O., Thiran, J.P.,2002. What and where in human audition: selective deficits following focal hemispheric lesions. Exp. Brain Res. 147, 8-15

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J., Hichwa, R.D.,2000. Subcortical and cortical brain activity during the feeling of self-generated emotions. Nat. Neurosci. 3, 1049-1056

Eggermont, J.J. and Ponton, C.W.,2002. The neurophysiology of auditory perception: from single units to evoked potentials. Audiol. Neurootol. 7, 71-99

Esslen, M., Pascual-Marqui, R.D., Hell, D., Kochi, K., Lehmann, D.,2004. Brain areas and time course of emotional processing. Neuroimage. 21, 1189-1203

Fletcher, P.C., Frith, C.D., Baker, S.C., Shallice, T., Frackowiak, R.S., Dolan, R.J.,1995. The mind's eye--precuneus activation in memory-related imagery. Neuroimage. 2, 195-200

Friederici, A.D., Mecklinger, A., Spencer, K.M., Steinhauer, K., Donchin, E.,2001. Syntactic parsing preferences and their on-line revisions: a spatio-temporal analysis of event-related brain potentials. Brain Res. Cogn Brain Res. 11, 305-323

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., Schlaug, G.,2003. Functional anatomy of pitch memory--an fMRI study with sparse temporal sampling. Neuroimage. 19, 1417-1426

Gaab, N. and Schlaug, G.,2003. The effect of musicianship on pitch memory in performance matched groups. NeuroReport 14, 2291-2295

Galaburda, A. and Sanides, F.,1980. Cytoarchitectonic organization of the human auditory cortex. J. Comp Neurol. 190, 597-610

Gottselig, J.M., Brandeis, D., Hofer-Tinguely, G., Borbely, A.A., Achermann, P.,2004. Human central auditory plasticity associated with tone sequence learning. Learn. Mem. 11, 162-171

Goydke, K.N., Altenmuller, E., Moller, J., Munte, T.F.,2004. Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. Brain Res. Cogn Brain Res. 21, 351-359

Grady, C.L., Van Meter, J.W., Maisog, J.M., Pietrini, P., Krasuski, J., Rauschecker, J.P.,1997. Attention-related modulation of activity in primary and secondary auditory cortex. Neuroreport 8, 2511-2516

Griffiths, T.D.,2001. The neural processing of complex sounds. Ann. N. Y. Acad. Sci. 930, 133-142

Griffiths, T.D. and Warren, J.D.,2004. What is an auditory object? Nat. Rev. Neurosci. 5, 887-892

Griffiths, T.D. and Warren, J.D.,2002. The planum temporale as a computational hub. Trends Neurosci. 25, 348-353

Griffiths, T.D., Warren, J.D., Scott, S.K., Nelken, I., King, A.J.,2004. Cortical processing of complex sound: a way forward? Trends Neurosci. 27, 181-185

Halpern, A.R., Zatorre, R.J., Bouffard, M., Johnson, J.A.,2004. Behavioral and neural correlates of perceived and imagined musical timbre. Neuropsychologia 42, 1281-1292

Hickok, G. and Poeppel, D.,2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67-99

Hyde, M.,1997. The N1 response and its applications. Audiol. Neurootol. 2, 281-307

Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B., Bharucha, J.J.,2002. The cortical topography of tonal structures underlying Western music. Science 298, 2167-2170

Jancke, L.,1996. The hand performance test with a modified time limit instruction enables the examination of hand performance asymmetries in adults. Percept. Mot. Skills 82, 735-738

Jancke, L., Buchanan, T.W., Lutz, K., Shah, N.J.,2001. Focused and nonfocused attention in verbal and emotional dichotic listening: an FMRI study. Brain Lang 78, 349-363

Jancke, L., Schlaug, G., Huang, Y., Steinmetz, H.,1994. Asymmetry of the planum parietale. Neuroreport 5, 1161-1163

Jancke, L. and Steinmetz, H.,1993. Auditory lateralization and planum temporale asymmetry. NeuroReport 5, 169-172

Jancke, L., Wustenberg, T., Scheich, H., Heinze, H.J.,2002. Phonetic perception and the temporal cortex. Neuroimage. 15, 733-746

Jones, S.J., Longe, O., Vaz, P.M.,1998. Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of 'streaming'? Electroencephalogr. Clin. Neurophysiol. 108, 131-142

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J.,2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163-178

Kaas, J.H. and Hackett, T.A.,1998. Subdivisions of auditory cortex and levels of processing in primates. Audiol. Neurootol. 3, 73-85

Kaas, J.H. and Hackett, T.A.,2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. U. S. A 97, 11793-11799

Kaas, J.H., Hackett, T.A., Tramo, M.J.,1999. Auditory processing in primate cerebral cortex. Curr. Opin. Neurobiol. 9, 164-170

Knosche, T.R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O.W., Friederici, A.D.,2005. Perception of phrase structure in music. Hum. Brain Mapp. 24, 259-273

Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lutkenhoner, B.,2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb. Cortex 13, 765-772

Langner, G., Sams, M., Heil, P., Schulze, H.,1997. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. J. Comp Physiol [A] 181, 665-676

Laufer, I. and Pratt, H.,2005. The 'F-complex' and MMN tap different aspects of deviance. Clin. Neurophysiol. 116, 336-352

Lewis, J.W., Wightman, F.L., Brefczynski, J.A., Phinney, R.E., Binder, J.R., DeYoe, E.A.,2004. Human brain regions involved in recognizing environmental sounds. Cereb. Cortex 14, 1008-1021

Liegeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., Chauvel, P.,1998. Contribution of different cortical areas in the temporal lobes to music processing. Brain 121, 1853-1867

Lutkenhoner, B. and Steinstrater, O.,1998. High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiol. Neurootol. 3, 191-213

Maeder, P.P., Meuli, R.A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J.P., Pittet, A., Clarke, S.,2001. Distinct pathways involved in sound recognition and localization: a human fMRIstudy. Neuroimage. 14, 802-816

Marcar, V.L. and Loenneker, T.,2004. The BOLD response: a new look at an old riddle. NeuroReport 15, 1997-2000

McAdams S and Bigand E, 1993. Recognition of sound sources and events. In: Thinking in Sound: The Cognitive Psychology of Human Audition. Oxford University Press, Oxford, pp 146-198

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., Krimphoff, J.,1995. Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. Psychol. Res. 58, 177-192

Menning, H., Roberts, L.E., Pantev, C.,2000. Plastic changes in the auditory cortex induced by intensive frequency discrimination training. NeuroReport 11, 817-822 Menon, V., Levitin, D.J., Smith, B.K., Lembke, A., Krasnow, B.D., Glazer, D., Glover, G.H.,

McAdams, S.,2002. Neural correlates of timbre change in harmonic sounds. Neuroimage. 17, 1742-1754

Meyer, M., Alter, K., Friederici, A.D., Lohmann, G., von Cramon, D.Y.,2002. FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum. Brain Mapp. 17, 73-88

Meyer, M., Steinhauer, K., Alter, K., Friederici, A.D., von Cramon, D.Y.,2004. Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang 89, 277-289

Meyer, M., Zaehle, T., Gountouna, E.-V., Barron, A., Jancke, L., Turk, A.,in press. Spectrotemporal processing during speech perception involves left posterior auditory cortex. NeurReport

Meyer, M., Zysset, S., von Cramon, D.Y., Alter, K.,2005. Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. Brain Res. Cogn Brain Res. 24, 291-306

Milner B, 1962. Laterality effects in audition. In: Interhemispheric Relations and Cerebral Dominance. John Hopkins University Press, Baltimore, pp 173-201

Mulert, C., Jager, L., Propp, S., Karch, S., Stormann, S., Pogarell, O., Moller, H.J., Juckel, G., Hegerl, U.,2005. Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI. Neuroimage. 28, 49-58

Mulert, C., Jager, L., Schmitt, R., Bussfeld, P., Pogarell, O., Moller, H.J., Juckel, G., Hegerl, U.,2004a. Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. Neuroimage. 22, 83-94

Mulert, C., Juckel, G., Augustin, H., Hegerl, U.,2002. Comparison between the analysis of the loudness dependency of the auditory N1/P2 component with LORETA and dipole source analysis in the prediction of treatment response to the selective serotonin reuptake inhibitor citalopram in major depression. Clin. Neurophysiol. 113, 1566-1572

Mulert, C., Pogarell, O., Juckel, G., Rujescu, D., Giegling, I., Rupp, D., Mavrogiorgou, P., Bussfeld, P., Gallinat, J., Moller, H.J., Hegerl, U.,2004b. The neural basis of the P300 potential. Focus on the time-course of the underlying cortical generators. Eur. Arch. Psychiatry Clin. Neurosci. 254, 190-198

Murayama, J., Kashiwagi, T., Kashiwagi, A., Mimura, M.,2004. Impaired pitch production and preserved rhythm production in a right brain-damaged patient with amusia. Brain Cogn 56, 36-

Naatanen, R. and Picton, T.,1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375-425

Nichols, T.E. and Holmes, A.P.,2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1-25

Overy, K., Norton, A.C., Cronin, K.T., Gaab, N., Alsop, D.C., Winner, E., Schlaug, G.,2004. Imaging melody and rhythm processing in young children. NeuroReport 15, 1723-1726

Pantev, C., Roberts, L.E., Schulz, M. Engelien, A., Ross, B.,2001. Timbre-specific enhancement of auditory cortical representations in musicians. NeuroReport 12, 169-174

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M.,1998. Increased auditory cortical representation in musicians. Nature 392, 811-814

Pardo, P.J., Makela, J.P., Sams, M.,1999. Hemispheric differences in processing tone frequency and amplitude modulations. NeuroReport 10, 3081-3086

Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D.,2002. Functional imaging with lowresolution brain electromagnetic tomography (LORETA): a review. Methods Find. Exp. Clin. Pharmacol. 24 Suppl C, 91-95

Pascual-Marqui, R.D., Lehmann, D., Koenig, T., Kochi, K., Merlo, M.C., Hell, D., Koukkou, M.,1999. Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. Psychiatry Res. 90, 169-179

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D.,1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49-65

Peretz, I. and Coltheart, M.,2003. Modularity of music processing. Nat. Neurosci. 6, 688-691

Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I.,2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage. 16, 331-348

Phillips, C., Rugg, M.D., Friston, K.J.,2002a. Anatomically informed basis functions for EEG source localization: combining functional and anatomical constraints. Neuroimage. 16, 678-695

Phillips, C., Rugg, M.D., Fristont, K.J.,2002b. Systematic regularization of linear inverse solutions of the EEG source localization problem. Neuroimage. 17, 287-301

Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J.,1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neurootol. 4, 64-79

Platel, H., Price, C., Baron, J.C., Wise, R., Lambert, J., Frackowiak, R.S., Lechevalier, B., Eustache, F.,1997. The structural components of music perception. A functional anatomical study. Brain 120, 229-243

Pratt, H., Bleich, N., Mittelman, N.,2005. The composite N1 component to gaps in noise. Clin. Neurophysiol. 116, 2648-2663

Rauschecker, J.P.,1998. Cortical processing of complex sounds. Curr. Opin. Neurobiol. 8, 516-521

Rauschecker, J.P., Tian, B., Hauser, M.,1995. Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268, 111-114

Rauschecker, J.P., Tian, B., Pons, T., Mishkin, M.,1997. Serial and parallel processing in rhesus monkey auditory cortex. J. Comp Neurol. 382, 89-103

Robin, D.A., Tranel, D., Damasio, H.,1990. Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. Brain Lang 39, 539-555

Samson, S.,2003. Neuropsychological studies of musical timbre. Ann. N. Y. Acad. Sci. 999, 144-151

Samson, S. and Zatorre, R.J.,1988. Melodic and harmonic discrimination following unilateral cerebral excision. Brain Cogn 7, 348-360

Samson, S. and Zatorre, R.J.,1994. Contribution of the right temporal lobe to musical timbre discrimination. Neuropsychologia 32, 231-240

Scherg, M. and von Cramon, D.Y., 1986b. Psychoacoustic and electrophysiologic correlates of central hearing disorders in man. Eur. Arch. Psychiatry Neurol. Sci. 236, 56-60

Scherg, M. and von Cramon. D.Y., 1986a. Evoked dipole source potentials of the human auditory cortex. Electroencephalogr. Clin. Neurophysiol. 65, 344-360

Schulte, M., Knief, A., Seither-Preisler, A., Pantev, C.,2002. Different modes of pitch perception and learning-induced neuronal plasticity of the human auditory cortex. Neural Plast. 9, 161-175

Scott, S.K. and Johnsrude, I.S.,2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci. 26, 100-107

Seither-Preisler, A., Krumbholz, K., Lutkenhoner, B.,2003. Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? Audiol. Neurootol. 8, 322-337

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E.,2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J. Neurosci. 23, 5545-5552

Shahin, A., Roberts, L.E., Trainor, L.J.,2004. Enhancement of auditory cortical development by musical experience in children. NeuroReport 15, 1917-1921

Sinai, A. and Pratt, H.,2003. High-resolution time course of hemispheric dominance revealed by low-resolution electromagnetic tomography. Clin. Neurophysiol. 114, 1181-1188

Specht, K. and Reul, J.,2003. Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task. Neuroimage 20, 1944-1954

Talairach J and Tournoux P, 1988. Co-Planer Stereotaxic Atlas of the Human Brain. Georg Thieme Verlag, Stuttgart.

Tervaniemi, M., Winkler, I., Näätänen, R.,1997. Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. NeuroReport 8, 2571-2574

Toiviainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M., Näätänen, R.,1998. Timbre Similarity: Convergence of Neural, Behavioral and Computational Approaches. Music Perception 16, 223-241

Vaughan, H.G., Jr. and Ritter, W.,1970. The sources of auditory evoked responses recorded from the human scalp. Electroencephalogr. Clin. Neurophysiol. 28, 360-367

Vitacco, D., Brandeis, D., Pascual-Marqui, R., Martin, E.,2002. Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing. Hum. Brain Mapp. 17, 4-12

Warren, J.D., Jennings, A.R., Griffiths, T.D.,2005. Analysis of the spectral envelope of sounds by the human brain. Neuroimage. 24, 1052-1057

Warren, J.D., Zielinski, B.A., Green, G.G., Rauschecker, J.P., Griffiths, T.D.,2002. Perception of sound-source motion by the human brain. Neuron 34, 139-148

Wessinger, C.M., VanMeter, J., Tian, B., Van, L.J., Pekar, J., Rauschecker, J.P.,2001. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J. Cogn Neurosci. 13, 1-7

Westbury, C.F., Zatorre, R.J., Evans, A.C.,1999. Quantifying variability in the planum temporale: a probability map. Cereb. Cortex 9, 392-405

Wong, D., Pisoni, D.B., Learn, J., Gandour, J.T., Miyamoto, R.T., Hutchins, G.D.,2002. PET imaging of differential cortical activation by monaural speech and nonspeech stimuli. Hear. Res. 166, 9-23

Zaehle, T., Wustenberg, T., Meyer, M., Jancke, L.,2004. Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. Eur. J. Neurosci. 20, 2447-2456

Zatorre, R.J.,1985. Discrimination and recognition of tonal melodies after unilateral cerebral excisions. Neuropsychologia 23, 31-41

Zatorre, R.J., Belin, P., Penhune, V.B.,2002. Structure and function of auditory cortex: music and speech. Trends Cogn Sci. 6, 37-46

Zatorre, R.J. and Binder, J.R., 2000. Functional and Structural Imaging of the Human Auditory System. In: Toga A and Mazziotta JC (Eds.), Brain Mapping: The Systems. Academic Press, San Diego, pp 365-402

Zatorre, R.J., Evans, A.C., Meyer, E.,1994. Neural mechanisms underlying melodic perception and memory for pitch. J. Neurosci. 14, 1908-1919

Zatorre, R.J., Mondor, T.A., Evans, A.C.,1999. Auditory attention to space and frequency activates similar cerebral systems. Neuroimage. 10, 544-554

Zatorre, R.J. and Penhune, V.B.,2001. Spatial localization after excision of human auditory cortex. J. Neurosci. 21, 6321-6328

5 Second Study

The Neuroplastic Enhancement of Auditory Evoked Potentials in Musicians Reflects an Influence of Expertise but not Attention

SIMON BAUMANN, MARTIN MEYER & LUTZ JÄNCKE

Div. Neuropsychology, Institute for Psychology, University of Zurich

Submitted to *The Journal of Neuroscience*

5.1 Abstract

Instrumental tones and in some instances simple sine wave tones were shown to evoke stronger auditory evoked responses in musicians compared to non-musicians. This effect was taken as an example for plasticity in the auditory cortex elicited by training. Up to the present, however, it is unknown whether an enlarged cortical representation for (instrumental) tones or increased neuronal activity provoked by focused attention in musicians accounts for the reported difference. In an attempt to systematically investigate the influence of attention to the processing of various tones we compared auditory evoked potentials recorded from musicians and non-musicians. During the EEG recording the participants were involved in tasks requiring more or less specific attention to specific sound features such as pitch or timbre. Our results demonstrate that attention does not affect the differences observed between musicians and non-musicians at the auditory event-related N1 potential (AEP). We also show that attention effects cannot explain the reported stronger P2 components in musicians. Current density mapping of the two components further indicates that

the differences between musicians and non-musicians originate, at least partially, from the auditory cortex. These results support the view that the increased AEP on tones in musicians reflects an enlarged neuronal representation for specific sound features of these tones. Finally, we show novel electrophysiological data demonstrating that musicians differ from non-musicians specifically in early attention focusing to specific sound features.

5.2

Introduction

Increasing evidence demonstrates that the auditory cortex is susceptible to plastic alterations. Animal studies showed reorganization in the primary auditory cortex (AI) of adult cats after cochlea damage (Rajan et al., 1993) and frequency training in monkeys resulted in an increased neuronal representation of the trained frequency bands (Recanzone et al., 1993). A seminal study reporting plasticity effects in the human auditory cortex (Pantev et al., 1998) observed increased MEG responses to piano tones compared to sine wave tones for pianists but not for non-musicians. This observation was interpreted as an effect of long-term exposure to piano sounds. Meanwhile some of the original findings have been slightly revised. Several studies utilizing Electroencephalography (EEG) or Magnetoencephalography (MEG) provide evidence that instrumental tones elicit increased responses compared to sine wave tones in non-musicians, too (Lutkenhoner et al., 2006; Shahin et al., 2005);Meyer et al., in press). Nevertheless, the direct contrast of musicians to non-musicians has been shown to reveal several enhanced auditory evoked potential (AEP) components or components of its magnetic equivalent the auditory evoked field (AEF) (Kuriki et al., 2006; Shahin et al., 2003; Shahin et al., 2005), but see Lütkenhöner et al. (2006) for counterevidence. These effects mainly pertain to the AEP/AEF[13] negativity N1 and the positivity P2. Enhancement of these components is also observed after training to discriminate specific sound features (Menning et al., 2000; Tremblay et al., 2001).

Referring to the studies by Recanzone et al. (1993), higher N1 and P2 amplitudes in music experts are usually interpreted as evidence for an increased neuronal representation for (instrumental) tones. However, not only a higher number of stimuli sensitive neurons but also a shift of attention can result in increased auditory activation and finally in larger auditory component amplitudes (Hillyard et al., 1973). In order to understand how plasticity affects the functional organization of the human brain it is imperative to distinguish between transient attention effects and permanent alteration of the neuronal architecture of the cortex.

Most of the studies described above addressed this problem using visual distractors (e.g. silent cartoon movies) to turn the participants' attention away from the auditory

[13] For the readers convenience AEP/AEF is simplified to "Auditory Evoked Response" (AER) when both meanings are included in the present document.

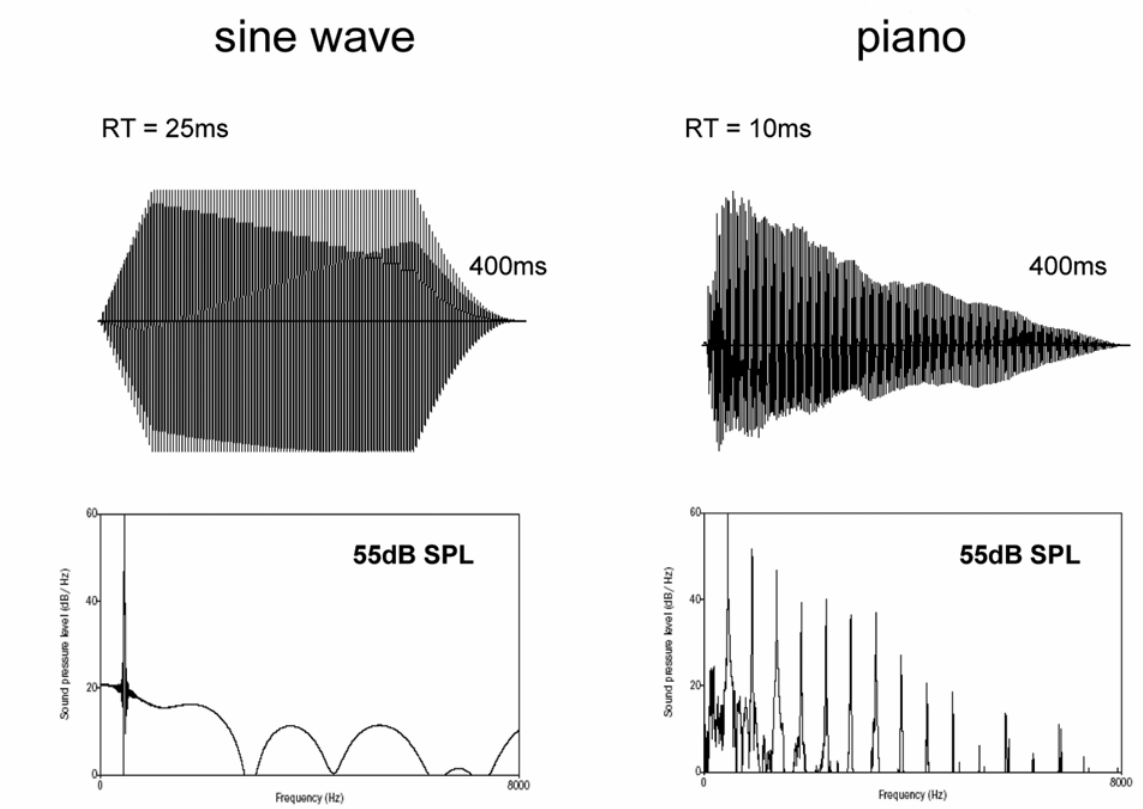
modality. However, it cannot be ruled out that musicians tend to direct their attention involuntarily to instrumental tones or to specific sound features such as pitch or timbre when it comes to sound perception because they have extensively been exposed to these sounds and acoustic features. Thus, we conducted this study to systematically investigate the influence of attention to music relevant sound features on the enhancement of the N1-P2 complex in musical experts. The primary aim is to address the following questions. First, for which AEP components are we able to replicate the reported difference between musicians and nonmusicians and which cortical areas may account for this difference? Second, is attention to specific sound features such as pitch or instrumental timbre systematically reflected by modulations of the AEP? And finally, if attention does modulate the AEP, to what extent does this effect differ between musicians and nonmusicians and could this effect explain the reported increase of the AEP in musicians.

5.3 Material and Methods

Twenty-six subjects with normal audiological status and no history of neurological pathology participated in this study. According to the Annet-Handedness-Questionnaire (AHQ) (Annett, 1970) and HDT (Steingruber1971; Jancke, 1996) all subjects were consistently right-handed. A group of 13 non-musicians (NM) subjects (5 female, 8 male, mean age ± SD of 26.2 ± 1.5 years) with no formal musical training and no history of musical instrument performance and a second group of 13 musicians (M) (4 female, 9 male, mean age ± SD of 25.6 ± 3.4 years) with a formal training starting at a mean age of 7.7 ± 2.7 years partake in this study. All but one of the musicians were either music students or members of an orchestra / jazz band and they practice their instruments daily between 1-5 hours. Most of the musicians play multiple instruments of which the most abundant were piano, violin, trumpet and saxophone. All of the subjects gave written consent in accordance with procedures approved by the local ethics committee and were paid for participation.

We presented a total of eight different tones using Hifi-headphones. The presented stimuli comprised short piano tones, trumpet tones, violin tones and sine wave tones at two different pitches (F4/349 Hz and B4/466 Hz, American notation). Details of the

stimuli are shown in Figure 1 and are described in (Meyer et al., in press). The stimuli were matched for intensity and the tones were presented at a sound pressure level of 55dB (double checked by applying a digital sound level meter[14]). Each of the four tone classes (sine wave, piano, trumpet and violin) was presented 80 times in one experimental block and 320 times in the entire experimental session. Each tone was made up of 40 F4 tones and 40 B4 tones. Thus, the total number of stimuli presented in the current study was 1280 (320 per block). Each tone was presented for 400 ms followed by a 1600 ms Inter-stimulus-Interval (ISI).



[14] Voltcraft 329, Conrad Electronics, Wernberg, Germany

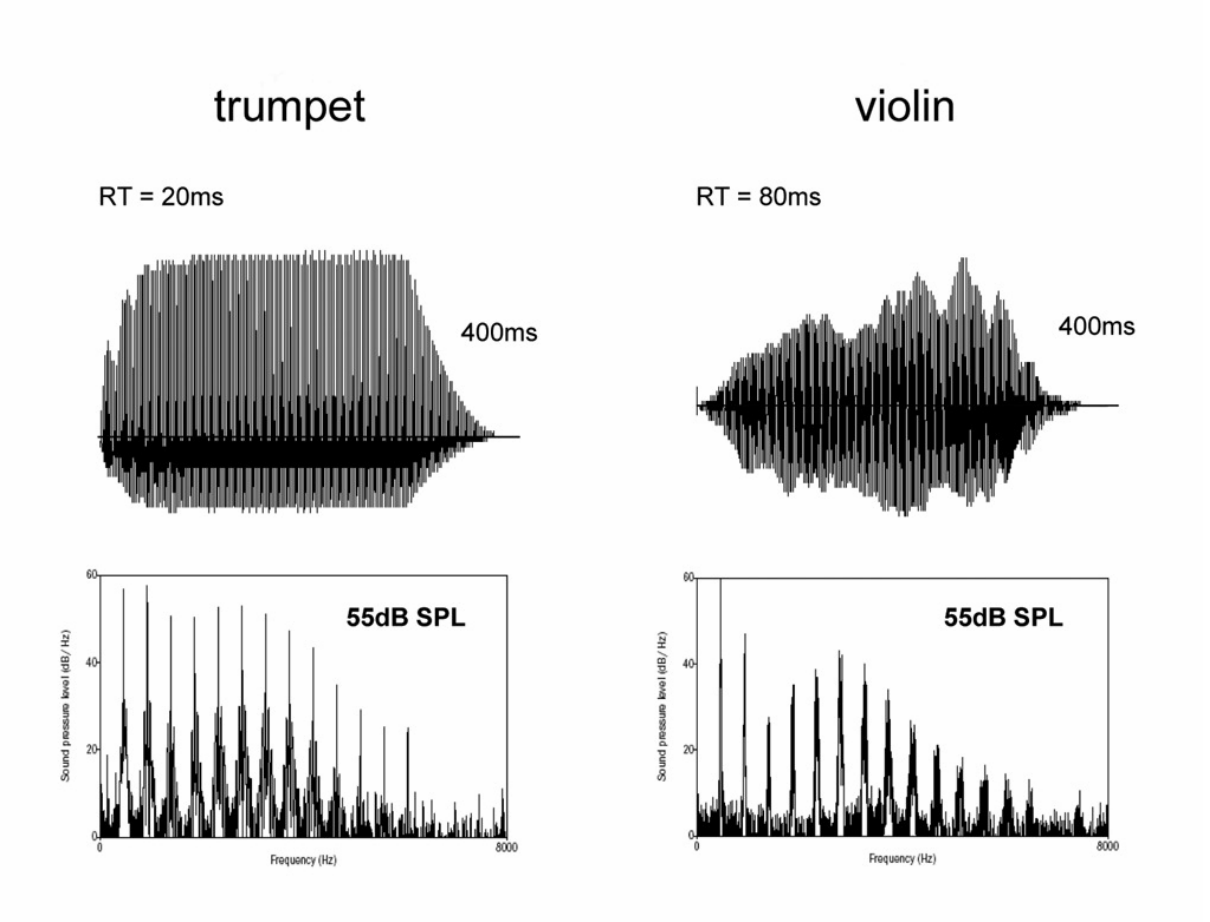


Fig. 1. *Stimulus sound envelopes and spectra.* Sound envelopes are displayed on the top panels the frequency spectra on the bottom panels.

Stimulation and recording of responses was controlled by the *Presentation* software (Neurobehavioral Systems, USA)[15]. Each session comprised four blocks with four different tasks. 1. undirected listening (control), 2. pitch discrimination, 3. target detection (sine wave tone), and 4. target detection (one out of three instrumental tones). During the control task the participants were required to press alternately one of two buttons with the right hand after presentation of each tone. During the other blocks participants were engaged in discrimination tasks. By means of the two buttons they had to indicate the pitch level or the absence or presence of either sine wave tones or one of the three instrumental tones. The order of the tasks/blocks was randomized.

The EEG (30 channels + 2 eye channels, subset of 10/10 system provided by Easy Cap) was recorded using an amplifier by Brainproducts (Germany). We applied

[15] http://nbs.neuro-bs.com

sintered Ag/AgCl electrodes and used the FCz position as reference. Electrode impedance was reduced to <10 kΩ by Electrogel conductant. The EEG data was sampled at 500 Hz and filtered with 100 Hz lowpass.

The raw data was further processed using Brain Vision Analyzer software (Brainproducts, Germany). After band pass filtering to 1-45 Hz, Independent Component Analysis (ICA) was applied in order to remove eye-movement- and in some cases muscle-artifacts (Jung et al., 2000). The processed data was re-referenced to an artificial reference derived from the average of all electrodes and epoched according to stimulus type (sine wave, piano, trumpet or violin) and task (as described above) into 700 ms segments, including a 200 ms prestimulus baseline. Epochs were averaged for each subject according to stimulus type and task collapsing over F4 and B4 pitches. In addition, trials comprising the three different instruments were collapsed into instrumental tone epoch averages. Thus, we ended up with 2 groups x 2 stimuli classes x 4 task conditions. Baseline correction relative to the -200 to 0 ms pre-stimulus time period was applied.

Auditory evoked potentials (AEPs) were identified at the Cz electrode site. Peaks of the N1 component were determined as the most negative potential between 50 and 150 ms after stimulus onset. Peaks of the P2 component were determined as the most positive component. Statistical analysis was performed by means of the SPSS software[16]. AEP peak values for N1 and P2 were evaluated by repeated-measures ANOVA analyzing the effects of GROUP[17](M, NM) x TONE(sine wave tones, instrumental tones) x TASK(control task, pitch detection task, sine wave tone detection task, instrumental tone detection task). Group differences of N1 and P2 peaks were compared for sine wave tones and instrumental tones by t-tests. In addition, AEP and global field power or root mean square (RMS) curves of the AEP between groups were compared by t-tests over all time points using the Brain Vision Analyzer software.

In order to examine attention to specific sound features we subtracted the potentials of the control task from the detection task potentials. Post-hoc analysis of these subtraction curves was performed at the time windows (TWs) between 150 and 200 ms and between 300 and 350 ms. The choice of these TWs is based upon a visual inspection of the difference curves in order to include the approximate peaks of the observed components. One-sample-t-tests were performed to test for attention

[16] http://www.spss.com

[17] GROUP was treated as “between subject” factor.

effects for specific sound features involving the two groups and the three detection tasks. In addition, the influence of GROUP(M, NM) and TASK(pitch detection, sine wave tone detection was analyzed by ANOVA. For all statistical tests an alpha level of 0.05 was applied and tests were corrected for sphericity violations where appropriate {Huynh1970}. All t-tests were calculated for two-tailed significance levels.

We applied an inverse linear solution approach – LORETA (low-resolution electromagnetic tomography[18] to estimate the neural sources of event-related scalp potentials (Pascual-Marqui et al., 1994; Pascual-Marqui, 1999; Pascual-Marqui et al., 2002). LORETA calculates the three dimensional distribution of electrically active neuronal generators in the brain as current density value (~μA/mm^2) based on the recorded scalp electric potentials. LORETA provides a solution for the *inverse problem* assuming that the smoothest of all possible activity distributions is the most plausible one to explain the data ("smoothness assumption"). Here, we determined the current density distribution for epochs of electrical brain activity on a dense grid of 2394 voxels at 7-mm spatial resolution. LORETA refers to a three-shell spherical model registered to the Talairach human brain atlas {Talairach1988}. The source locations were therefore provided as (x, y, z) coordinates in the stereotactic Talairach space.

In order to verify the localization of the N1 and the P2 component we calculated the LORETA current density maximum (~μA/mm^2) from the grand mean over all subjects of each group at the TWs (latency 100 ± 20 ms) for the N1 and (latency 200 ± 20 ms) for the P2 component, respectively. The same procedure was applied for the localization estimation of the processes underlying the observed components at the post-hoc defined time periods between 150-200 ms and 300-350 ms respectively. We estimated the spatial localization for cortical areas showing enhanced activity in musicians vs. controls for sine wave tones and instrumental tones at the N1 TW (latency 100) by performing voxel-wise t-tests on the individual current density maps. Multiple comparison correction was performed by applying a nonparametric randomization test (Nichols & Holmes, 2002). Effects of the musician vs. non-musician comparison over the primary and secondary auditory cortex were evaluated by conducting a regions of interest (ROI) analysis based on individual mean current density values at ROIs defined by Brodmann Areas 41/42 (exact location and size of the ROIs is described in Meyer et al. (in press). Current density values were

[18] http://www.unizh.ch/keyinst/NewLORETA/LORETA01.htm

extracted applying the LORETA ROI extractor tool[19] and then subjected to t-tests checking for group differences.

5.4 Results

Behavioral data

The performance of the musicians in the tasks was close to perfect. The rate of correct button presses was 99.5 ± 0.2 (control), 99.2 ± 0.4 (pitch discr.), 99.3 ± 0.5 (sine wave discr.) and 99.3 ± 0.3 (instr. tone discr.). Nonmusicians performed particularly in the pitch discrimination task worse: 99.2 ± 1.0 (control), 89.8 ± 9.2 (pitch discr.), 96.4 ± 5.6 (sine wave discr.) and 98.1 ± 1.15 (instr. tone discr.). T-tests revealed significant group differences for pitch discrimination (T_{24} = 3.68; p = 0.003) and for instrumental tone discrimination. The differences for sine tone discrimination was marginally not significant (T_{24} = 1.85; p = 0.089).

AEP amplitudes

Figure 2 shows separate AEP (a, c) and RMS (b, d) curves for musicians and nonmusicians and for sine wave tones and instrumental tones. All four tasks are collapsed in one curve for each condition and group. Clear peaks of N1 and P2 amplitudes were observed in all conditions at about 100 ms and 200 ms after sound onset, respectively. We observed higher mean AEP amplitudes of the N1 component (significant between 44-150 ms for sine wave tones and 72-118 ms for instrumental tones) and higher mean RMS potentials for the N1 (significant between 76-130 ms for sine wave tones and 72-104 ms for instrumental tones) and P2 time points for musicians. A main effect for GROUP was observed for N1 ($F_{(1, 24)}$ = 11.9; p = 0.002)

[19] www.novatecheeg.com/Downloads.html

but not for P2 peaks, indicating that increased potentials for musicians versus non-musicians were only significant for the N1 component. Further main effects were observed for TONE for N1 ($F_{(1, 24)}$ = 44.2; p < 0.001) and P2 peaks ($F_{(1, 24)}$ = 70.7; p < 0.001), and for TASK only for the P2 peaks ($F_{(1, 24)}$ = 4.8; p=0.01). Of all possible interactions only the one for TASK x GROUP for the P2 peaks was significant ($F_{(1, 24)}$ = 3.4; p = 0.037).

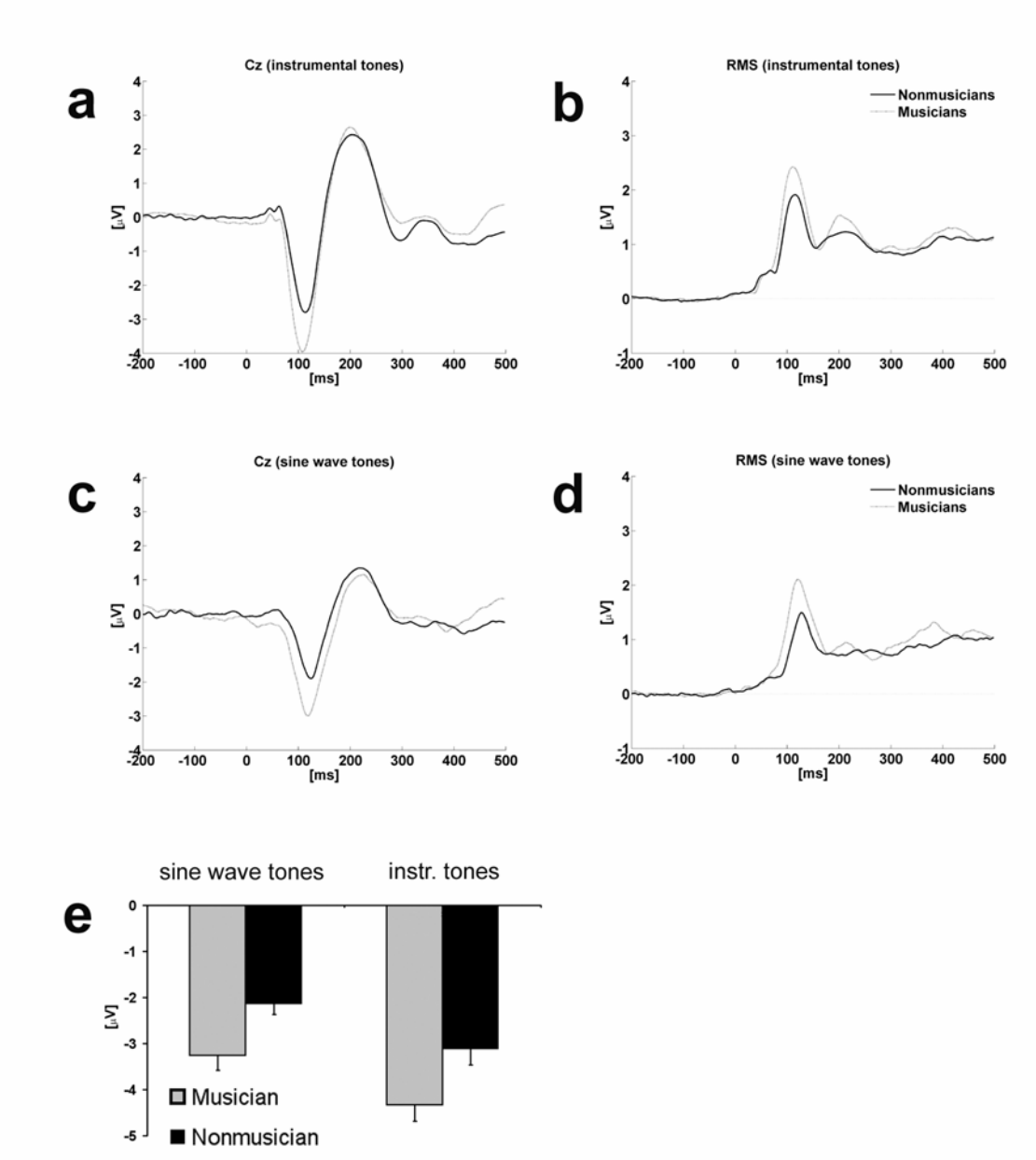


Fig. 2. *AEP and RMS curves.* AEP curves recorded at the Cz electrode site are shown for instrumental tones (**a**) and sine wave tones (**c**) respectively. Curves are displayed in light grey for musicians and in black for nonmusicians. **b**, **d** The corresponding root mean square (or global field power) values are displayed in the panels on the right of the AEPs. Sound onset is set to time point zero. Panel **e** shows the results of the N1 peak analysis

Effects of sound feature attention

Figure 2e shows the results for t-tests comparing the AEP amplitude peaks of the N1 and P2 components for sine wave tones and instrumental tones between the two groups (pooling over all tasks). We observed increased N1 potentials for musicians independent of the tone class (T_{24} = 2.75; p = 0.012 / sine wave tones, T_{24} = 2.43; p = 0.023 / instrumental tones), but no significantly different P2 potentials. The difference between sine wave tones and the corresponding instrumental tones was significant in all cases (p < 0.001).

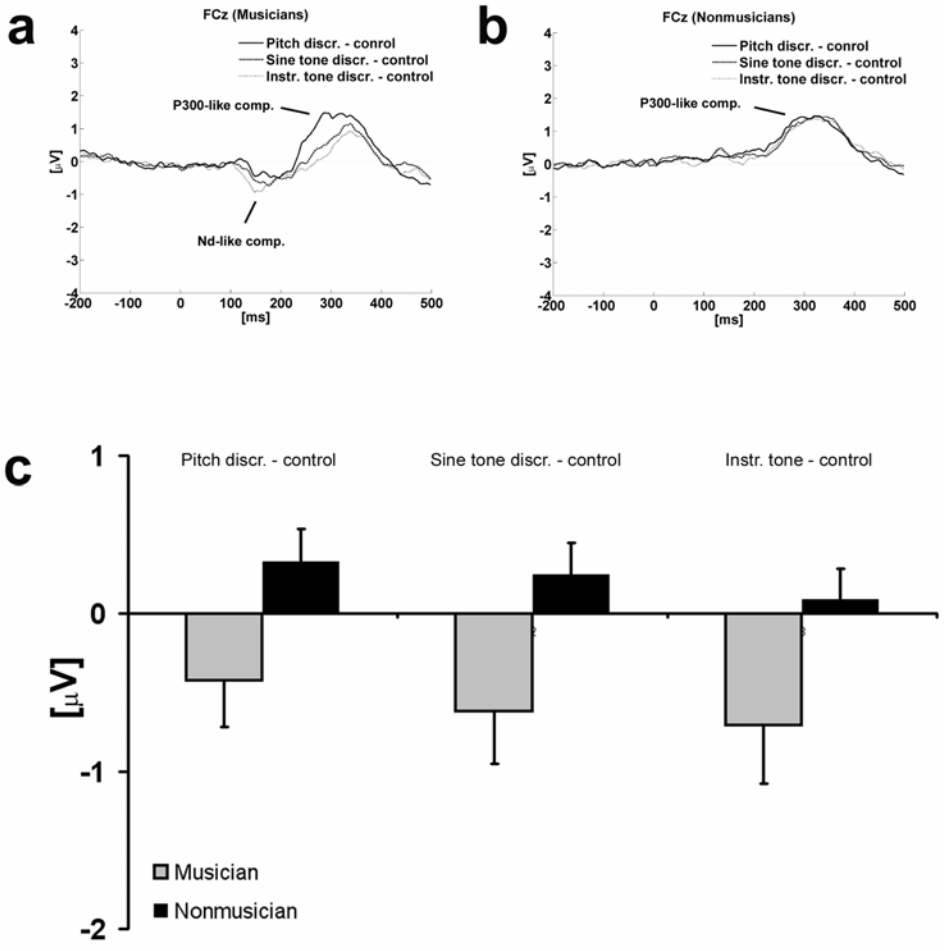


Fig. 3. *Difference potential curves for specific vs. unspecific attention.* Difference curves of subtracted potentials at FCz for specific attention vs. unspecific attention (control) are displayed in the top panels for musicians (**a**) and nonmusicians (**b**). Pitch discrimination vs. control is drawn in black, instrumental tone discrimination vs. control is drawn in grey and the sine wave tone discrimination curve is represented by a dotted line. (**c**) mean potential values for all difference curves are shown for the time window 150-200 ms.

In order to investigate the influence of attention to specific sound features (pitch, timbre) we calculated the difference potentials between the auditory feature detection tasks versus the control task, which required unspecific auditory attention. Because we observed no TASK x TONE interactions in the previous peak analysis we collapsed the two tone classes into one difference potential for the three task comparisons. The difference potentials are shown in Figure 3 with separate panels for musicians (a) and non-musicians (b). The strongest task effects emerged 200 ms after sound onset peaking in positivities between 300-350 ms latency. These effects were observed in all “detection task vs. control task difference curves” and in both subject groups. The positivities were not only distinct from the classical AEPs N1 and P2 in their temporal occurrence but also in their spatial pattern. The maximal difference potentials were consistently at the FCz electrode site slightly anterior to Cz. In musicians, we observed also a negativity of the described difference potentials between 100 and 250 ms after sound onset with maxima between 150 ms and 200 ms, which was not existent in nonmusicians. A t-test between the detection tasks and the control task at the TW of 150-200 ms and 300-350 ms respectively (Fig. 3c) revealed that the later effect was clearly significant for the non-musicians ($T_{12} = 4.81$; $p < 0.001$ / pitch detection, $T_{12} = 5.09$; $p < 0.001$ / sine wave tone detection, $T_{12} = 3.77$; $p = 0.003$ / instrumental tone detection) and for all but one detection tasks in the musicians ($T_{12} = 4.68$; $p = 0.001$ / pitch detection, $T_{12} = 2.56$; $p = 0.025$ / sine wave tone detection). The comparison of the instrumental tone detection task with the control task marginally missed significance ($T_{12} = 1.73$; $p = 0.11$). Likewise, the comparisons at the 150-200 ms latency TW did not reach significance in both groups although the effect was close to significance in the musicians. However, an ANOVA for the task effects at the same TWs using the variables GROUP and TASK(pitch detection, sine wave tone detection, instr. tone detection) revealed a group effect for the TW 150-200 ms ($F_{(1, ?)} = 4.6$; $p = 0.042$) but not for the later TW 300-350 ms after sound onset. Neither of the two TWs showed significant task effects between the detection tasks. In summary, we find common attention effects in both groups around 300 ms after sound onset but the musicians show clear differences from the non-musicians in earlier attention around 150 ms latency.

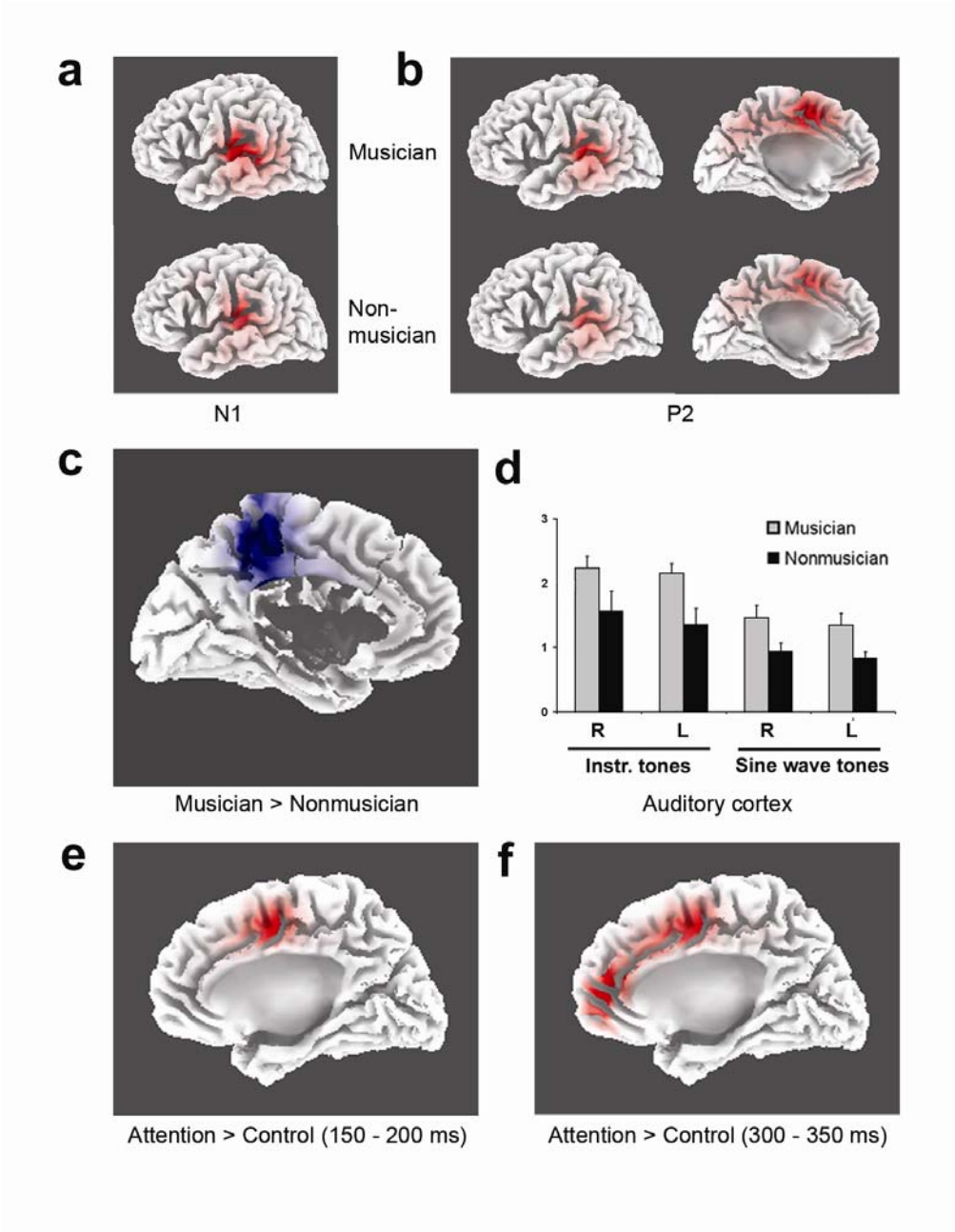


Fig. 4. *Current density distributions.* The maxima of the current density distributions for instrumental tones at the time windows for N1 (80-120 ms), (**a**) and P2 (180-220), (**b**) are displayed for musicians (top panels) and nonmusicians (bottom panels). Panel **c** shows the significant voxels ($p<0.05$, corrected for mult. comp.) derived from voxel-wise comparison between musicians and nonmusicians for instrumental tones. The data from a Region of Interest Analysis of the same contrast in the auditory cortex is shown in the panel on the right (**d**). Significant differences are marked with asterisks. In the bottom panels the current density maxima of the difference waves derived from the specific vs. unspecific attention tasks are displayed at the time windows 150-200 ms (**e**) and 300-350 ms (**f**) covering the components that were most sensitive to attention. The maps derived from the instrumental discrimination vs. control conditions in musicians form representative examples for the data listed in more detail in Table 1.

Source estimation

Applying the LORETA source estimation, we observed bilateral maxima over the primary and secondary auditory cortices for the sine wave tones and the instrumental tones for both the N1 and the P2 TWs (Figure 4a, b). We observed a prominent secondary maximum in the anterior cingulum (ACC) at the P2 window. Talairach coordinates and current density values of all reported maxima are listed in Table 1.

Table I. *List of current density maxima.* Current density maxima are listend by calculated time window including current density value [µA/mm$_2$]*10$_{-3}$, talairach coordinates, hempisphere and cortical area defined by Brodmann area (BA). For Voxel-wise comparisons the t-value is reported. The threshold for significant voxels corrected for multiple comparisons is at a t-value of 3.21 ($p<0.05$).

Time window	Condition	Current density	Coordinates			Hemisph.	Cortical area
			x	y	z		BA
N1 (80-120 ms)	Sine Tone Mus.	1.72	-59	-32	8	L	42
		1.77	60	-39	-15	R	22
	Instr. Tone Mus.	2.01	-59	-32	8	L	42
		2.13	60	-39	8	R	22
	Sine Tone Nonmus.	1.4	-59	-32	8	L	42
		1.44	60	-32	-15	R	42
	Instr. Tone Nonmus.	1.61	-59	-32	8	L	42
		1.58	60	-39	15	R	42
P2 (180-220 ms)	Sine Tone Mus.	1.38	-59	-39	8	L	22
		1.41	60	-39	15	R	42
		1.23	4	3	43	R	24
	Instr. Tone Mus.	1.56	-59	-39	8	L	22
		1.5	60	-39	15	R	22
		1.32	-4	-4	50	L	24
	Sine Tone Nonmus.	1.27	-59	32	8	L	42
		1.18	60	-39	22	R	22
		1.28	4	3	43	R	24
	Instr. Tone Nonmus.	1.41	-59	-39	8	L	22
		1.42	60	-39	15	R	42
		1.3	4	3	43	R	24

150-200 ms	Musicians						
	Instr. Discr. - Cont.	0.81	4	-4	50	R	24
	Sine Discr. - Cont.	0.7	4	-4	50	R	24
	Pitch Discr. - Cont	0.56	-3	11	50	L	24
300-350 ms	Nonmusicians						
	Instr. Discr. - Cont.	1.07	4	10	43	R	32
	Sine Discr. - Cont.	1.02	4	3	43	R	24
	Pitch Discr. - Cont	1.05	4	3	43	R	24
	Musicians						
	Instr. Discr. - Cont.	0.9	4	-4	43	R	24
		0.67	-3	45	-6	L	10
	Sine Discr. - Cont.	0.84	4	-4	43	R	24
		0.78	4	45	1	R	32
	Pitch Discr. - Cont	1.1	4	-4	50	R	24
Voxel-wise comp	Mus. vs. Nonmus.	t - value					
N1 (80-120 ms)	Instr. Tone	3.34	-10	-53	57	L	7
	Sine Tone	3.11	-10	-46	57	L	7

The comparison between the estimated current density values for musicians and non-musician controls revealed significant higher current density (corrected for multiple comparisons) in the right superior parietal lobe (BA5/7) (Figure 4c).

We observed a maximum in the same area for the sine wave tones but this was mariginally not significant on a corrected level (Table 1). A closer investigation of the current density distribution in the auditory cortex by a ROI analysis showed significantly higher current densities for the musicians against controls for instrumental tones (T_{24} = 2.14; p = 0.045 / left, T_{24} = 2.70; p = 0.014 / right) and sine

wave tones (T_{24} = 2.23; p = 0.037 / left, T_{24} = 2.48; p = 0.023 / right) in both the left and the right hemisphere (Figure 4d).

The current density maxima for the observed components of the task difference waves at the TWs between 150-200 ms latency and 300-350 ms latency are listed in Table 1. (See also Figure 4e, f for examples). The maxima at the 150-200 TW in the musicians were observed in the anterior cingulum for all tasks. This was also the case for the maxima at the 300-350 TW for the two groups. However, in the musicians prominent secondary maxima were observed in the posterior cingulum and the subcallosal area

5.5 Discussion

First, we will briefly reconsider to what extent our data is in line with previous investigations showing an enhanced AER to instrumental and sine wave tones in musicians compared to controls. We will discuss the effect of selective attention to specific sound features on the AER in the two groups and finally we will point out the evidence against the hypothesis that increased selective attention is the reason for stronger AER in musicians.

AEP differences between musicians and controls

Our data clearly supports previous studies showing an enhanced N1 component of the AER in musicians compared to non-musicians. This is particularly remarkable because in accordance with numerous previous studies our data show a considerable intersubject variability (see also Lütkenhöner et al. (2006)). This variability might be a reason why we do not find a significantly increased P2 component in musicians versus nonmusicians although we find mean RMS potential increases of comparable magnitude as for N1. Another source of variance that affects specifically the P2 time range is discussed in the following section. Due to the high variance and because several studies showed an enhanced AER P2 component

(Kuriki et al., 2006; Shahin et al., 2003; Shahin et al., 2005) we carefully avoid to speculate about a particular role for one of the two main components (N1 and P2) of the AER in musical expertise.

In contrast to earlier reports (Pantev et al., 1998; Pantev and Lutkenhoner, 2000) we find enhanced N1 amplitudes in musicians versus non-musicians for sine wave tones in addition to the group effect for instrumental tones. This is well in line with the data presented by Shahin et al. (2003; 2005) who reported also magnified P2 and the N1c component of the AER obtained from similar contrasts. The observation of similar effects in the two groups for the N1 component for instrumental and sine wave tones is confirmed by the fact that we find no GROUP x TONE interaction. This indicates that also not specifically trained tones evoke additional activity in musicians compared to non-musicians. In order to gain more information on the nature of this additional activity we estimated the location of the underlying cortical generators by applying the LORETA approach to map the current density distribution.

We observed bilateral current density maxima in the primary or secondary auditory cortex for all classes of tones for the N1 and P2 component. This reasonable finding is in agreement with previous dipole source estimations (Hari et al., 1987; Pantev et al., 1996) and LORETA estimations (Meyer et al., in press). According to this result the strongest neuronal responses in the N1 and P2 time range are generated in the auditory cortex. In order to test the idea that additional neuronal generators result in enlarged representation of tones in the musician's auditory cortex, a voxel-wise comparison of the current density distribution derived from both groups was performed. The current density based ROI analysis revealed significantly increased activity in the primary and secondary auditory cortices of musicians compared to controls which indicates that enlarged neuronal representation might be an explanation for the observed stronger AERs. However, the location of the maxima of the current density comparison also suggest that additional cortical areas outside the auditory cortex should be considered as potentially accommodating neural generators contributing to the group difference observed on the N1 component. Early associative crossmodal activity within the first 100 ms after stimulus onset evoked by auditory and visual stimuli was previously demonstrated by intracortical recordings (Schroeder and Foxe, 2002). Although we can not make firm conclusions about the role of the posterior parietal cortex (BA5/7) in our task, several fMRI studies demonstrated a privileged role of this brain area in musicians vs. nonmusicians

comparisons (Hasegawa et al., 2004; Haslinger et al., 2005), music training (Stewart et al., 2003)Stewart2005} and imagery of music performance (Meister et al., 2004). Involvement of this cortical area is explained by audio-visual and visuo-motor transformation processes which are essential for music performance. This would also explain why we find the strongest activity in the parietal cortex when the participants listen to instrumental tones. Interestingly, in an other fMRI study on musical performance, we also find the strongest differences in brain responses between pianists and non-musicians in the posterior parietal cortex (BA7) while they listen to piano pieces (S. Baumann, unpublished observations). Despite additional activation outside the auditory cortex we conclude that our present data strongly supports the notion that the auditory cortex shows higher activity in musicians compared to non-musicians while listening to (instrumental) tones.

Influence of attention on the AEP

Whereas the ANOVA analysis for the N1 peak potentials did not reveal task effects and therefore rules out an influence of attention to musically relevant sound features at this stage, the same analysis for the P2 potentials showed a different picture. Not only did we find a main effect of task at this component but the GROUP x TASK interaction indicated that this effect differed between musicians and non-musicians. This observation was confirmed by inspection of the potential difference curves between the sound feature detection tasks and the control task which required unspecific auditory attention. In the TW between 150-200 ms after sound onset, the time range of the P2 component, we observed negative components for all specific attention tasks versus the control task exclusively for the musicians. Although the negativity at this TW marginally fails to reach significance at the FCz electrode, the difference between the groups is significant. Furthermore, we observed a later component peaking at a latency of about 300 ms, a component that was detected in both groups and all three detecting vs. control difference curves. A complex composed of an earlier negativity followed by a later positivity in this time range is known from attention modulation in “oddball”-paradigms and it is called N2/P3 complex. A role of attention for the reported components is supported by the estimated location of their current density maxima. The ACC has repeatedly been

proposed to play an important part in the network for attention control (Hopfinger et al., 2000; Posner and Dehaene, 1994; Turken and Swick, 1999).

A particular obvious central negativity between 150-200 ms is reported in the literature for difference curves between attended and ignored stimuli and is referred to as Nd (negative displacement) (Hansen and Hillyard, 1980) and explained by channel selection (Naatanen, 1982). Because we observed a similar component in a contrast of specific sound feature attention to general auditory attention in our study we would call it Nd-like component. A previous study by Münte et al. (Munte et al., 2001; Nager et al., 2003) already demonstrated modulation of the Nd in the context of attention and musical expertise. Specific attention to sounds from spatially different sources evoked stronger Nd components in conductors compared to pianist and nonmusicians, which was explained by their acquired skill to locate specific instruments in a concert hall. The Nd component was seen as an index for the initial selection of stimuli according to certain stimulus features and therefore its enhancement in conductors was taken as an example for neuroplasticity due to training. In our study we can similarly explain the Nd-like component by an intensive training to the attended stimuli features pitch and musical timbre and thus demonstrate a novel effect of plasticity on top down processes in skilled musicians.

Although we can show that attention to specific musically relevant sound features can affect the AEP, none of the demonstrated effects can account for the enhanced AEP in musicians vs. controls. Whereas the N1 component is not modulated by sound feature attention the demonstrated effect at a latency of 300 ms is too late to explain the reported AER differences between musicians and non-musicians. The only attention effect during the relevant time range that is actually shown to differ between the groups is the reported Nd–like component. Enhanced attention, however, results in an increasing Nd negativity at the central electrodes. Thus, increased attention in musicians results in a reduced AEP positivity P2, which is measured at the same electrodes and in the same time range. Hence, increased attention to specific sound features in musicians compared to nonmusicans is not an explanation for the demonstrated group differences of the AEP. In contrary, attention effects resulting in the Nd component would even reduce the amplitude of the P2 or at least increase the variability at this stage of processing. This could be a further explanation for the unclear results at the P2 processing stage in our data.

5.6 Conclusion

The present study reveals clear evidence for enhanced activity in the auditory cortex of musicians compared to nonmusician controls. A detailed investigation on the influence of sound feature specific attention demonstrates that the observed effects do not cause the observed differences between musicians and non-musicians. Therefore, our results support the view that an enlarged neuronal representation of tones or of specific sound features of tones accounts for the observed group effects rather than a difference in attention levels. Finally, we revealed a difference between musicians and nonmusicians in the neurophysiological response to attended tones which suggests the existence of additional early processes in musicians for specific sound feature detection.

Aknowledgments:

During the preparation of this manuscript S.B. and M.M. were supported by the Swiss National Science Foundation (SNF) 46234101. The authors thank Stefan Elmer and Nathalia Estévez for assistance in recording the data.

5.7 References

Annett, M.,1970. A classification of hand preference by association analysis. Br. J. Psychol. 61, 303-321

Hansen, J.C. and Hillyard, S.A.,1980. Endogenous Brain Potentials Associated with Selective Auditory Attention. Electroencephalography and Clinical Neurophysiology 49, 277-290

Hari, R., Pelizzone, M., Makela, J.P., Hallstrom, J., Leinonen, L., Lounasmaa, O.V.,1987. Neuromagnetic Responses of the Human Auditory-Cortex to Onsets and Offsets of Noise Bursts. Audiology 26, 31-43

Hasegawa, T., Matsuki, K.I., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., Sadato, N.,2004. Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. Cognitive Brain Research 20, 510-518

Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H., Ceballos-Baumann, A.O.,2005. Transmodal sensorimotor networks during action observation in professional pianists. Journal of Cognitive Neuroscience 17, 282-293

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W.,1973. Electrical Signs of Selective Attention in Human Brain. Science 182, 171-180

Hopfinger, J.B., Buonocore, M.H., Mangun, G.R.,2000. The neural mechanisms of top-down attentional control. Nature Neuroscience 3, 284-291

Jancke, L.,1996. The hand performance test with a modified time limit instruction enables the examination of hand performance asymmetries in adults. Percept. Mot. Skills 82, 735-738

Huynh H., Feldt L.S. (1970). Conditions under which the mean square rations in repeated measurements designs have exact F distributions. Journal of the American Statisctical Association 65, 1582-1589

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J.,2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163-178

Kuriki, S., Kanda, S., Hirata, Y.,2006. Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. Journal of Neuroscience 26, 4046-4053

Lutkenhoner, B., Seither-Preisler, A., Seither, S.,2006. Piano tones evoke stronger magnetic fields than pure tones or noise, both in musicians and non-musicians. Neuroimage 30, 927-937

Meister, I.G., Krings, T., Foltys, H., Boroojerdi, B., Muller, M., Topper, R., Thron, A.,2004. Playing piano in the mind - an fMRI study on music imagery and performance in pianists. Cognitive Brain Research 19, 219-228

Menning, H., Roberts, L.E., Pantev, C.,2000. Plastic changes in the auditory cortex induced by intensive frequency discrimination training. Neuroreport 11, 817-822

Meyer, M., Baumann, S., Jäncke, L., 2006. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. Neuroimage (in press)

Munte, T.F., Kohlmetz, C., Nager, W., Altenmuller, E.,2001. Neuroperception - Superior auditory spatial tuning in conductors. Nature 409, 580-580

Naatanen, R.,1982. Processing Negativity - An Evoked-Potential Reflection of Selective Attention. Psychological Bulletin 92, 605-640

Nager, W., Kohlmetz, C., Altenmuller, E., Rodriguez-Fornells, A., Munte, T.F.,2003. The fate of sounds in conductors' brains: an ERP study. Cognitive Brain Research 17, 83-93

Nichols, T.E. and Holmes, A.P.,2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1-25

Pantev, C. and Lutkenhoner, B.,2000. Magnetoencephalographic studies of functional organization and plasticity of the human auditory cortex. Journal of Clinical Neurophysiology 17, 130-142

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M.,1998. Increased auditory cortical representation in musicians. Nature 392, 811-814

Pantev, C., Roberts, L.E., Elbert, T., Ross, B., Wienbruch, C.,1996. Tonotopic organization of the sources of human auditory steady-state responses. Hearing Research 101, 62-74

Pascual-Marqui, R.D.,1999. Reviews of methods for solving the EEG inverse problem. International Journal of Bioelectromagnetism 1, 75-86

Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D.,2002. Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. Methods Find. Exp. Clin. Pharmacol. 24 Suppl C, 91-95

Pascual-Marqui, R.D., Michel, C.M., Lehmann, D.,1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49-65

Posner, M.I. and Dehaene, S.,1994. Attentional Networks. Trends in Neurosciences 17, 75-79

Rajan, R., Irvine, D.R.F., Wise, L.Z., Heil, P.,1993. Effect of Unilateral Partial Cochlear Lesions in Adult Cats on the Representation of Lesioned and Unlesioned Cochleas in Primary Auditory-Cortex. Journal of Comparative Neurology 338, 17-49

Recanzone, G.H., Schreiner, C.E., Merzenich, M.M.,1993. Plasticity in the Frequency Representation of Primary Auditory-Cortex Following Discrimination-Training in Adult Owl Monkeys. Journal of Neuroscience 13, 87-103

Schroeder, C.E. and Foxe, J.J.,2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Cognitive Brain Research 14, 187-198

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E.,2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. Journal of Neuroscience 23, 5545-5552

Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B.,2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport 16, 1781-1785

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., Frith, U.,2003. Brain changes after learning to read and play music. Neuroimage 20, 71-83

Tremblay, K., Kraus, N., Mcgee, T., Ponton, C., Otis, B.,2001. Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. Ear and Hearing 22, 79-90

Turken, A.U. and Swick, D.,1999. Response selection in the human anterior cingulate cortex. Nature Neuroscience 2, 920-924

6 Third Study

A Network for Audio-Motor Coordination in Skilled Pianists and Non-Musicians

SIMON BAUMANN[1], SUSAN KOENEKE[1], CONNY F. SCHMIDT[1, 2], MARTIN MEYER[1], KAI LUTZ[1] & LUTZ JANCKE[1]

[1] Div. Neuropsychology, Institute for Psychology, University of Zurich
[2] Institute for Biomedical Engineering, University of Zurich and ETH Zurich

Revised for *Cerebral Cortex*

Short version published in *Ann. N. Y. Acad. Sci. (2005). 1060, 186-88*

6.1 Abstract

Playing a musical instrument requires efficient auditory and motor processing. Fast feed forward and feedback connections that link the acoustic target to the corresponding motor programs need to be established during years of practice. The aim of our study is to provide a detailed description of cortical structures that participate in this audio-motor coordination network in professional pianists and non-musicians. In order to map these interacting areas using functional magnetic resonance imaging (fMRI), we considered cortical areas that are concurrently activated during covert piano performance and motionless listening to piano sound. Furthermore we investigated to what extent interactions between the auditory and the motor modality happen involuntarily. We observed a network of predominantly secondary and higher order areas belonging to the auditory and motor modality. The

extent of activity was clearly increased by imagination of the absent modality. However, this network did neither comprise primary auditory nor primary motor areas in any condition. Activity in the lateral dorsal premotor cortex (PMd) and the pre-supplementary motor cortex (preSMA) was significantly increased for pianists. Our data imply an intermodal transformation network of auditory and motor areas which is subject to a certain degree of plasticity by means of intensive training.

6.2 Introduction

From several points of view high level music performance is an interesting topic to study the neural underpinnings of action and perception and to address the question to what extent the human cerebral cortex is modified structurally and functionally due to training. On one hand, playing a musical instrument poses great demands on the human motor system. Complex movement coordination is required at tremendous speed and accuracy. On the other hand, there are probably no professionals that challenge musicians in their expertise in the auditory domain. Music performance requires detection of minimal changes in pitch and rhythm to ensure a perfect artistic outcome. Thus, highly skilled musicians are an ideal model to investigate function and plasticity of the auditory and the motor cortex (Munte et al., 2002; Schlaug, 2001). The combination of proficiency in both the auditory and the motor modality, however, makes musicians particularly interesting for the studying of interaction and coordination between both modalities. The Information flow between sensory processing and motor planning areas is crucial in music performance (Janata and Grafton, 2003). Fast feedforward and feedback connections are required to coordinate auditory input and motor output (Bangert and Altenmuller, 2003). Furthermore, these connections depend on cortical processing circuits which are capable of transforming auditory information into a code that is appropriate for use by the motor system and vice versa. An impressive amount of data about visuo-motor transformation processes in humans and monkeys has been published (Caminiti et al., 1991; Ellermann et al., 1998; Grefkes et al., 2004; Rizzolatti et al., 1996), reviewed by Burnod et al. (1999b). A preeminent structure consistently appearing in these studies is the premotor cortex. However, little is known about comparable audio-motor transformation centers so far (Hickok et al., 2003; Hickok and Poeppel, 2000).

Most of the previously published studies on music performance in professionals or amateurs focused either on music perception or instrument playing. Recent studies investigated brain activation on musical motor performance (Lotze et al., 2003; Meister et al., 2004). As expected, a wide range of primary, secondary (premotor cortex (PM), supplementary motor area (SMA)) and other motor and somatosensory related areas (e.g. basal ganglia, cerebellum) were involved. Interestingly, Lotze et al. (2003) also demonstrated differences between amateurs and professionals in the

primary and secondary auditory cortices. Although some studies had been set out to study combined auditory and motor control processes (Kristeva et al., 2003; Parsons et al., 2005), only very few of them focused on the interaction between both modalities. To our knowledge there is only one study which consequently addresses this issue (Bangert et al., 2001; Bangert and Altenmuller, 2003). In this study DC-EEG scalp maps recorded from piano novices during an audio-motor task become increasingly similar to the maps of professional pianists during the course of an audio-motor training. Haueisen & Knösche (2001) and Popescu et al. (2004) used a different approach to address audio-motor interaction. These two MEG studies observed activation in motor areas during pure auditory music stimulation in musicians (the former) or in non-musicians (the latter). The results imply either direct or indirect connectivity between auditory and motor areas.

The entirety of the aforementioned studies which investigate audio-motor-interactions are based on electrophysiological recordings. This guarantees high temporal resolution which is an advantage for the discovery of fast interactions between the modalities of interest. However, imaging technologies like PET or fMRI benefit from higher spatial resolution. Increased spatial resolution may be crucial to distinguish between primary and secondary sensory and motor structures. Furthermore, precise localization information helps to decide whether involuntary motor activity in musicians is indeed related to purposeful finger movement rather than unspecific motor activity (e.g., the urge of Jazz listeners to tap the rhythm of the music with their limbs). The tendency to react to music and rhythmic sound by tapping, drumming or even dancing is known from many cultures. However, to our knowledge there is no systematic neurophysiological study on involuntary motor activity induced by music which is not specific to a certain body part.

In a different context, Hickok et al. (2003) carried out an interesting analysis to examine an audio-motor network. In order to reveal cortical areas involved in perception and production during speech and singing, fMRI-data were recorded during a passive listening task and during an active auditory rehearsal task. Only brain areas which were active in both tasks were considered to be audio-motor integration areas. Hickok et al. (2003) reported bilateral premotor cortex (PM), bilateral inferior frontal gyrus (IFG) and an area in the posterior part of the Sylvian fissure in the left hemisphere. Additionally, the left superior temporal sulcus (STS) supported the speech task while the music task recruited the contralateral STS.

The value of studying music and in particular highly trained musicians as a model for cortical plasticity has been increasingly recognized in the last decade. Several anatomical and functional studies examined the effect of plastic changes due to intensive musical training. Increased grey matter volume for musicians was observed in the anterior corpus callosum (Schlaug et al., 1995), the cerebellum (Hutchinson et al., 2003; Schlaug, 2001), and in primary sensorimotor and premotor areas (Gaser and Schlaug, 2003). All these brain regions play an important role in either fine motor control or bimanual information transfer, both vital processes for music performance. An influential MEG study (Elbert et al., 1995) showed that fingers of the left hand of violinists show stronger representation by means of signal amplitude in the primary somato-sensory cortex than those of a control group. The same method was used to reveal increased auditory cortical representations for musical timbre in violinists and trumpeters relative to non-musicians (Pantev et al., 2001). fMRI studies, however, showed rather a decrease of intensity of activation in motor areas in musicians versus non-musicians. Further studies observed weaker activity in primary and secondary motor areas (primary motor area (M1), supplementary motor area (SMA), preSMA and cingulate motor area (CMA)) for pianists compared to non-musicians in either bimanual (Jancke et al., 2000b) or unimanual piano playing like tapping tasks (Hund-Georgiadis and von Cramon, 1999). Even the comparison of professional vs. amateur violinists during the performance of a Mozart concerto by Lotze et al. (2003) revealed mainly stronger activity in motor areas for the amateurs. The lower activity in cortical motor control areas has commonly been attributed to diminished neural effort needed for a particular motor task performance.

In contrast to previously accomplished studies, the main interest of the present study is crossmodal activity, i.e. responses in the motor cortex elicited by auditory stimulation and activity in the auditory cortex triggered by piano (finger) movements. We hypothesize that these transmodal activities are stronger in pianists compared to non-musicians. This should especially be evident for music stimuli, an assumption based on the idea that musicians might have stronger connections between auditory and motor areas which results in an efficient translation of finger actions into auditory music representations and vice versa.

Another outstanding question is whether transmodal information transfer occurs completely involuntary or by the use of voluntary imagery of the missing modality.

The aforementioned studies by Haueisen & Knösche (2001) and Popescu et al. (2004) imply that information transfer from the auditory to the motor system has at least an involuntary component. However, various studies demonstrated that imagery of sound is sufficient to activate areas in the auditory cortex (Bunzeck et al., 2005; Halpern and Zatorre, 1999; Janata, 2001; Zatorre et al., 1996) and that imagery of music performance leads to activation in motor areas (Kristeva et al., 2003; Lotze et al., 2003; Meister et al., 2004).

The aim of our study is to provide a detailed description of cortical and subcortical areas involved in an audio-motor integration network using piano performance as an example. In order to find areas communicating with each other during piano performance, we specifically search for areas which are coincidentally active during both pure motor and pure auditory tasks. Furthermore, we investigate to what extent this network is reliant on voluntary, top down control and to what extent decades of training (as normally pertaining to musicians) change the activity pattern of the network. According to these aims we propose three hypotheses: 1.) Professional musicians show strong activation in a distributed network specialized to process audio-motor associations. 2.) Hearing musical pieces will automatically (implicitly) evoke activations in motor areas of musicians. Vice versa, playing musical pieces without hearing the tune will automatically evoke activations in auditory areas. 3) Pianists have stronger transmodal representation for music stimuli compared to non-musicians.

6.3 Materials and Methods

Subjects

7 highly skilled pianists from the Zurich Conservatory of Music (6 female; mean age: 25.7 ± 3.2 years) and 7 control subjects (3 female; mean age: 30.0 ± 5.0) took part in the study. The pianists received at least ten years of extensive piano training (age of commencement: 5-13) and the time of exercise comes to 3-5 h per day. Control subjects were matched according to the education level of the musicians. They did not receive any special musical training in their past life. All subjects were consistent right-handers, as determined by the Annett-Handedness-Questionnaire (AHQ)

(Annett, 1970). Furthermore, the hand dominance test (HDT) (Jancke, 1996; Steingruber, 1971) was used as a measure for manual lateralization. The experiments conform to the declaration of Helsinki and were approved by the local ethics committee.

Task design

Subjects had to perform different auditory and motor tasks while being placed in the MR scanner. During the pure auditory tasks subjects were instructed to listen to either a piece of music (Mozart, Sonata Facile) or scales in c-major continuously played up and down. Importantly, the instruction precluded overt finger movements. The two groups of subjects were familiar to the Mozart piece and the sound of scales respectively, and were able to reproduce the melody vocally. Stimulus material consisted of 15 s parts of the Mozart piece converted from compact disc to wave files (44.1 kHz, stereo) and digitally recorded piano tones (44.1 kHz, stereo) arranged to scales of equal duration. During the pure motor tasks subjects were asked to play either a piece of music (Mozart, Sonata Facile) or 5-tone c-major scales without any auditory feedback. All finger movements were carried out on a plastic board. Since non-musicians were evidently not able to accomplish fingerings adequate to the musical piece, they were instructed to carry out internally-paced, complex, bimanual finger movements. Furthermore, prior to experimental sessions, the group of non-musicians was informed on how to carry out the fingerings of a c-major scale. The advantage of two different tasks is that we can collect data from a task for which the motor performance is as similar as possible (scales) and from another task which is closer related to music (Mozart piece). The participants were instructed to choose a performance speed according to the audio presentation.

We are well aware that the use of a plastic board is a compromise between feasibility and authenticity. Not even the keyboard of an electronic piano can exactly mimic the specific key stroke characteristics of a mechanical piano. In the case of this study, however, we are mainly interested in the crossmodal and mental representation of piano performance, which should not be affected by the specific key stroke characteristics.

The motor tasks were carried out either with the right hand, left hand, or bimanually. This partition was designed as a control task with respect to an upcoming study and it

will not be a subject of discussion in the context of this publication. The four task conditions (1) "Listen to Mozart", (2) "Listen to Scales", (3) "Play Mozart" (Play Complex fingerings for non-musicians), and (4) "Play Scales" were carried out in two modes. In the *VOLUNTARY* mode subjects were instructed to explicitly imagine the absent modality during listening or playing (motor or auditory imagination). In the *INVOLUNTARY* mode subjects were told to focus on a particular distractor task while listening or playing. The distractor task required participants to count changes in brightness of a small grey square presented in the middle of the screen throughout the whole experimental run. This task was to be performed during ON and OFF blocks. The distracting stimulus was also present during tasks in *VOLUNTARY* mode, but there was no need to attend to it.

Thus, each subject took part in a total of 8 experimental conditions that were distributed across two separate sessions. Typically, subjects performed the 4 auditory conditions ("Listen to Mozart", "Listen to Scales" in either *VOLUNTARY* or *INVOLUNTARY* mode) on one day and the 4 motor conditions ("Play Mozart", "Play Scales" in either *VOLUNTARY* or *INVOLUNTARY* mode) on the other. The order of the two experimental sessions as well as the order of the 4 runs within a session was pseudo-randomized across subjects in order to control for order effects.

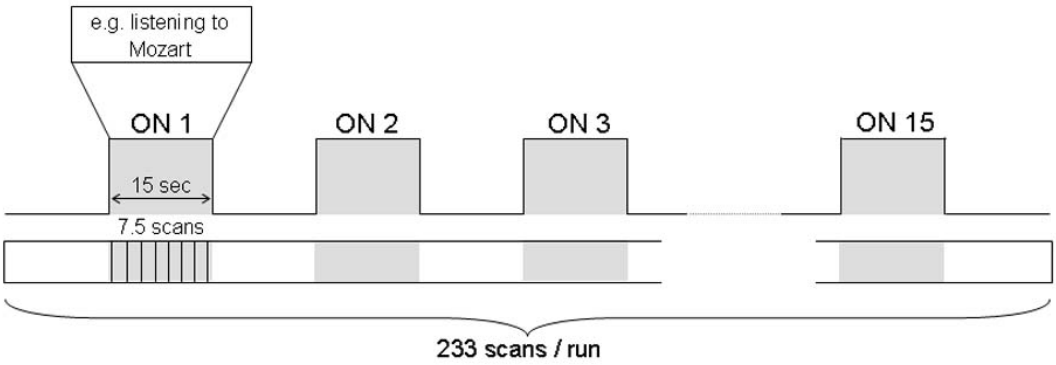


Fig. 1. Scan design with 15 ON blocks and 16 OFF blocks.

The tasks were run with a classical fMRI box-car design alternating rest (15 s) and activation (15 s) blocks. One experimental run consisted of 16 resting (OFF) and 15 activation (ON) blocks (Fig. 1). During the activation (ON) blocks the auditory and motor tasks were presented. During the OFF-blocks, subjects were instructed to either passively watch the grey square in the middle of the screen or counting the

number of its color changes, depending on the current assignment of a particular run to either the *VOLUNTARY* or INVOLUNTARY mode. For each of the 8 experimental conditions one session was conducted. Thus, the task condition did not change during the course of one run, however, the presented stimuli did. With regard to the auditory conditions different 15-sec-samples of the Mozart sonata were presented. The scale stimuli did not differ from one another. Visual presentation during auditory conditions was limited to the small grey square of the distractor task. During the motor tasks two hands were presented, one on each side of the square (subjects viewed the back of the hands). A green-colored circle that was placed around the left, the right or both hands indicated what hand(s) to use for a given 15-sec ON-block. The stimuli were presented via a video-projector onto a translucent screen which subjects viewed inside the scanner via a mirror. Prior to every session subjects got visual instruction regarding the task characteristics of the upcoming run.

Scanning procedure

Functional magnetic resonance imaging was performed on a Philips Intera 3T whole-body MR unit equipped with a transmit–receive body coil and a commercial eight-element head coil array (MRI Devices Corporation, Waukesha WI, USA). Three-dimensional anatomical images of the whole brain were obtained by using a T1-weighted three-dimensional, spoiled, gradient echo pulse sequence (TR = 20 ms, TE = 2.30 ms, flip angle = 20°, FOV = 220×220, acquisition matrix = 224×224, voxel size = 0.98mm × 0.98mm × 0.75 mm^3, 180 slices, slice thickness 0.75mm).
Functional data were obtained from 28 transverse slices covering the whole brain using a single-shot EPI technique with SENSE R=2.0 (Pruessmann et al., 1999) (TR = 2000 ms, TE = 35 ms, flip angle = 75°, FOV = 220×220, acquisition matrix = 80×80, voxel size = 2.75mm × 2.75mm × 4.5 mm^3). In order to reduce the acoustic noise factor a particular "SofTone Mode" was applied (de Zwart et al., 2002).

Data analysis

Image analysis was performed on a PC using MATLAB 6.5 (Mathworks Inc., Natiek, MA, USA) and SPM2 (http://fil.ion.ucl.ac.uk/spm). All images were realigned to the first volume, corrected for motion artifacts, normalized (2mm × 2mm × 2 mm) into

standard stereotaxic space (EPI-template provided by the Montreal Neurological Institute), and smoothed using an 8mm full-width-at-half-maximum Gaussian kernel. Activated voxels were identified by the “General Linear Model” approach. At the first level of analysis a statistical model for each subject was computed, applying a box-car model, convolved with a standard hemodynamic response set and eliminating low-frequency noise. A statistical parametric map of the T-statistic was generated for each voxel to test hypotheses about regionally specific condition effects. Linear contrasts were employed for each subject and condition, as suggested by Friston et al. (1995). The resulting set of voxel values for each contrast yields a statistical parametric map of the T-statistic [SPM(*T*)]. The following two contrasts have been calculated for each subject: "pure audio (all conditions) > OFF", "pure motor (all conditions) > OFF". Additionally these contrasts have been broken down for *VOLUNTARY* vs. *INVOLUNTARY* mode resulting in "pure audio (*VOLUNTARY* conditions) > OFF", "pure audio (*INVOLUNTARY* conditions) > OFF_{invol}", "pure motor (*VOLUNTARY* conditions) > OFF", "pure motor (*INVOLUNTARY* conditions) > OFF_{invol}". For the group comparison “pianists vs. non-musicians” each of these four conditions was split into a “MUSIC” and a “SCALES” subcondition, receiving 8 baseline-contrasts in total. Because the MUSIC and the SCALES subcondtions showed basically the same pattern and we observed no significant difference in the transmodal activations of interest between these tasks, we decided to pool the data in order to simplify the results.

To infer from our observations to population effects, a “second-level-analysis” was employed in order to compare the conditions across the entire group using the “contrast-images” obtained from the individual subjects. All t-maps resulting from the group analysis were thresholded using a height-threshold of $p = .001$ (uncorrected for multiple comparisons). Subtreshold activity was re-checked on a p = .005 (uncorrected) level.

In order to statistically verify the existence of concurrently activated regions, that means neuronal regions that are activated during both the pure auditory and the pure motor tasks, the minimum of both group contrasts (t-maps) was calculated. An additional spatial extent threshold of $n = 7$ voxel was applied to the resulting map of audio-motor activations. Furthermore, in order to investigate to what extent interactions between the auditory and the motor modality happen involuntarily, the same procedure was applied for *VOLUNTARY* and *INVOLUNTARY* conditions

separately. Between-group comparisons for our proposed audio-motor network were masked with the areas that had been concurrently activated in the pure motor and pure auditory condition. Threshold for t-maps was set to p = 0.05 (corrected for volume of mask). Only transmodal activity was considered.

In order to quantify activity patterns in distinct auditory and motor areas further analyses were performed using a region-of-interest (ROI) approach. Anatomical ROIs were created using MRIcro 1.4 (www.mricro.com) based on anatomical definitions of the functional areas as defined in earlier studies (Fink et al., 1997; Jancke et al., 1994; Roland and Zilles, 1996; Schmahmann et al., 2000; Steinmetz et al., 1990; Steinmetz et al., 1996). The following ROIs were used for both hemispheres: M1, lateral premotor (PM), preSMA, SMA proper, the cerebellar cortices, and the cerebellar nuclei, Heschl's Gyrus (HG), Planum Temporale (PT) and Planum Polare (PP), (Fig. 3C). For each ROI mean T-values were computed for each subject and subjected to further statistical analysis using SPSS 11.5 (SPSS Inc. 1989 – 1999; http://spss.com). According to the experimental paradigm a three-factor, repeated-measures analyses of variance (ANOVA) with factors "hemisphere" (left vs. right) "task" (auditory vs. motor) and "imagination mode" (VOLUNTARY vs. INVOLUNTARY) as independent variables was calculated for each ROI. Post-hoc analyses were carried out using Bonferoni-corrected t-tests for paired samples applying the correction procedure proposed by Holm (1979).

6.4 Results

fMRI measures

Executed movements vs. OFF

In general, we observed motor activations in the group analyses similar to previous studies which have mapped musical instrument performance (Lotze et al., 2003; Meister et al., 2004; Parsons et al., 2005). Peak activations were observed in the primary motor and somatosensory cortices with the maximum at x = 40 y = -38 z = 54 (MNI coordinates). Further activations comprised bilateral secondary sensorimotor areas such as PMd, preSMA/SMA, inferior parietal lobule and motor

related areas such as PMv, basal ganglia (lentiform nucleus) and cerebellum. The activations in the PMv extended seamlessly into the opercular part of the IFG. In addition to these sensory-motor areas we observed activations in the STG anterior and posterior to the Heschl's gyrus and in the parietal operculum in both hemispheres. The spread of motor induced activations is displayed in Fig. 2B and Tab. 1.

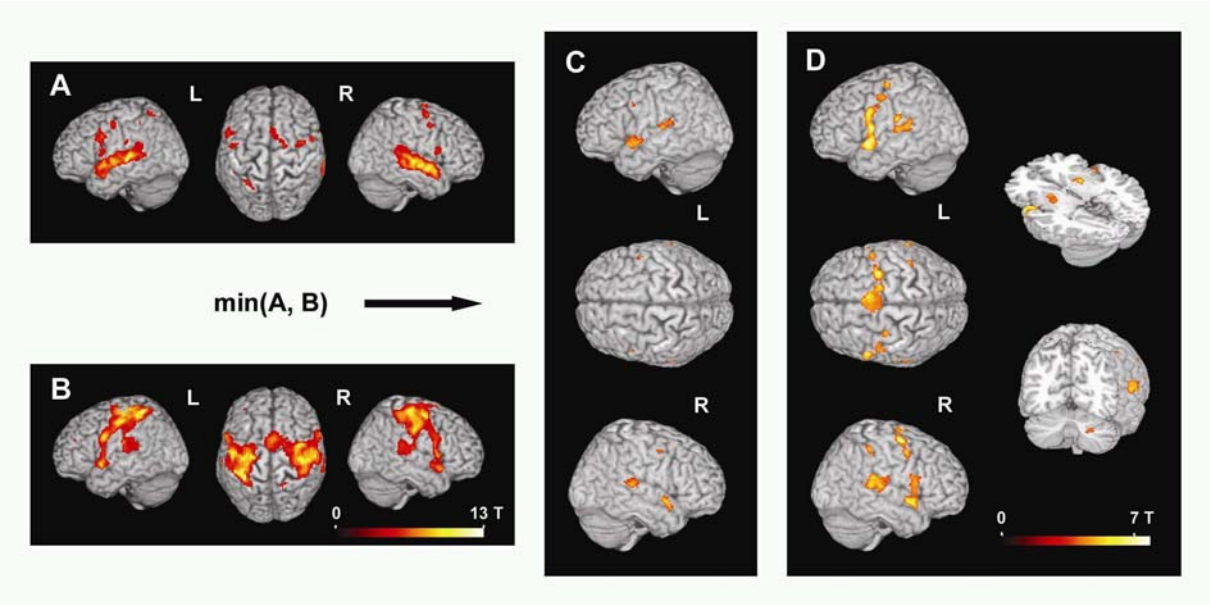


Fig. 2. Activations for all PURE AUDITORY vs. OFF conditions (p<0.001, uncorrected) (A). Activations for all PURE MOTOR vs. OFF conditions (p<0.001, uncorrected) (B). Scaling of T-values for A and B is indicated by the scale-bar in B. Minimum of all PURE AUDITORY and PURE MOTOR conditions in the INVOLUTARY mode (p<0.001, uncorrected) (C). Minimum of all PURE AUDITORY and PURE MOTOR conditions in the VOLUTARY mode (p<0.001, uncorrected) (D). Scaling of T-values for A and B is indicated by the scale-bar in D.

Sound listening vs. OFF

This contrast reveals bilateral activity in the group analyses in the entire STG with a maximum in the middle section of the STG at MNI coordinates x = -50, y = -12, z = 6 (Fig. 2A, Tab. 1). In addition to these well known auditory areas we observed activity in the ventral and dorsal premotor corex, the SMA, and the IFG (pars opercularis) in the two hemispheres. Lateralized activity was present in the cerebellum and the basal ganglia on the left and in two parietal foci of activation (parietal operculum, inferior parietal lobule) on the right hemisphere. Activation in the cerebellum was bilateral in all but 3 subjects. However, due to poor overlap only one cluster of

activated voxels on the left side remained after applying the threshold in the group analysis.

Table 1: Maxima of activated clusters expressed in MNI space for the motor and auditory conditions compared to the baseline ($p < 0.001$, $k > 7$).

All motor/auditory conditions vs. OFF

Cluster		**MNI coordinates**				
k	**t-value**	**x**	**y**	**z**	**Hem.**	**Location**
Motor						
12650	12.95	-40	-38	54	L	post_cent
2878	11.56	-16	-54	-32	L	CB
531	10.15	24	-8	-8	R	NL
1115	10	-26	-8	-10	L	NL
269	8.21	14	-22	-10	R	thal
36	5.12	18	-54	66	R	SPL
10	4.45	8	22	30	R	pre_cent
15	4.27	-34	44	26	L	MFG
Auditory						
3197	12.42	50	-12	6	R	STG
3444	11.02	-58	-26	2	L	STG/STS
272	6.86	-36	-44	36	L	IPL
84	6.53	42	-2	60	R	MFG
44	6.51	54	4	48	R	pre_cent
49	6.25	22	0	-6	R	NL
567	5.99	16	-6	74	R	SMA
103	5.69	-32	-56	62	L	SPL
16	5.61	-42	16	22	L	IFG
97	5.33	58	12	12	R	IFG
73	5.23	-52	-8	48	L	pre_cent
38	4.57	24	-70	-34	R	CB
13	4	-24	-8	48	L	pre_cent

Cerebellum (CB), inferior frontal gyrus (IFG), inferior parietal lobe (IPL), medial frontal gyrus (MFG), nucleus lentiformis (NL), postcentral gyrus (post_cent), precentral gyrus (pre_cent), supplementary motor area (SMA), superior parietal lobe (SPL), superior temporal gyrus (STG), superior temporal sulcus (STS), thalamus (thal).

Combined playing & listening vs. OFF

The primary focus of the study was to depict the areas that were concurrently active in the playing and the listening conditions. These areas comprised mainly secondary sensory, motor and higher order cortex. As described in the methods part we carried out a separate group analysis for involuntary (Fig. 2C) and voluntary (Fig. 2D) involvement of the missing modality. We observed partially overlapping areas activated in both contrasts. The jointly activated areas comprised bilateral planum temporale (PT), planum polare (PP) and PMd. Further areas were significantly activated in the VOLUNTARY mode, namely, bilateral preSMA/SMA, PMv, IFG (pars opercularis) and basal ganglia (lentiform nucleus). Unilateral activations were observed in the cerebellum, the inferior parietal lobule and in the superior temporal sulcus (STS) in the right hemisphere and in the parietal operculum (SII, secondary somatosensory cortex) in the left hemisphere. For this condition, bilateral PMv, IFG (pars opercularis) and basal ganglia ranked among the most reliably activated structures, although these three cortical areas had not even survived a threshold of $p<0.05$ (uncorrected) in the INVOLUNTARY mode in contrast to the preSMA/SMA, the cerebellum and the STS. Locations of t-value maxima for activated clusters are summarized in Tab 2.

Maximum t-values and the number of significantly activated voxels were higher in all areas for the VOLUNTARY compared to the INVOLUNTARY mode in the proposed audio-motor network. This difference is supported by the ROI analysis. All regions except for the PP and the PT show significantly higher ($p<0.05$) mean t-values for the VOLUNTARY compared to the INVOLUNTARY contrast considering the modality that was not stimulated (Fig. 3A). There was no area that showed a stronger BOLD signal in the INVOLUNTARY mode.

Interestingly, we did not find any activation that was concurrently present in the auditory and motor conditions of the group analysis in motor-, somatosensory- or auditory primary areas. We detected weakly activated voxels in primary sensory and motor areas in single subjects. However in the group analysis, the minimum of activation for both pure auditory and motor stimulation showed a negative beta weight correlation in most portions of the primary cortices when compared with rest. Mean t-values of selected regions of interest are displayed in Fig. 3A. Primary areas like the HG and M1 show negative mean t-values that are well below the positive

Table 2: Maxima of significantly activated clusters expressed in MNI space for the contrasts comparing the motor and auditory condition vs. baseline ($p < 0.001$, $k > 7$). In addition, the difference between pianists vs. non-musicians is shown for the imagery condition ($p < 0.05$, corrected for small volume, $k > 7$).

Minimum(Motor vs. OFF, Auditory vs. OFF)

Cluster	**MNI coordinates**					
k	**t-value**	**x**	**y**	**z**	**Hem.**	**Location**
Involuntary						
91	5.3	58	8	-10	R	PP
147	5.09	-56	12	-10	L	PP
248	4.96	-46	-30	10	L	PT
141	4.83	64	-32	8	R	PT
11	4.27	54	0	48	R	PMd
9	4.13	-54	8	36	L	PMd
Voluntary						
563	7.75	-54	10	-6	L	PP
563	6.32	-56	8	32	L	PMv
1122	7.22	34	-4	60	R	PMd
1122	6.17	8	4	56	R	SMA/pSMA
187	6.27	24	0	-6	R	NL
465	5.98	48	8	-12	R	PP
412	5.57	64	-38	10	R	PT/STS
50	5.42	-60	-24	24	L	par_op
56	5.33	54	2	48	R	PMd
239	5.31	-20	8	-6	L	NL
122	5.15	-34	-14	62	L	PMd
289	5.1	-58	-26	12	L	PT
88	5.06	46	-40	50	R	IPL
72	4.6	-52	-6	48	L	PMd
44	4.59	22	-68	-34	R	CB
Pianists vs. Non-musicians (imagery condition, Mozart)						
28	5.98	34	-4	60	R	PMd
7	4.55	-2	6	46	L	pSMA

Cerebellum (CB), inferior parietal lobe (IPL), nucleus lentiformis (NL), parietal operculum (par_op), dorsal premotor cortex (PMd), ventral premotor cortex (PMv), planum polare (PP), planum temporale (PT), pre-supplementary motor cortex (preSMA), supplementary motor cortex (SMA), superior temporal sulcus (STS).

values of secondary areas like the PT, PP, PMd and preSMA. This is in striking contrast to the situation for areas directly related to the stimulated modality where mean t-values of the primary areas are on comparable or even higher levels than the secondary areas (Fig 3B).

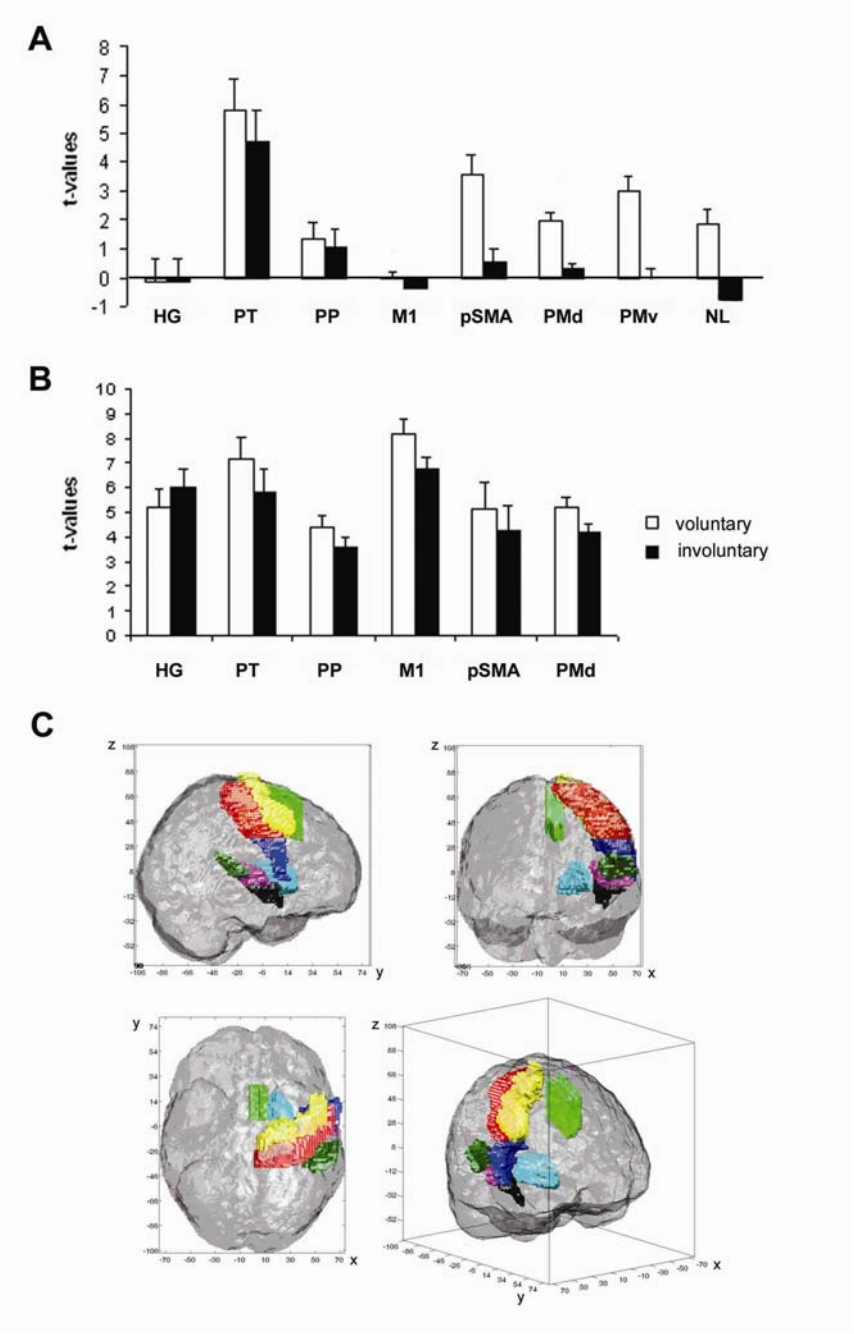


Fig. 3.
Mean T-values of the ROI analysis for the transmodally stimulated cortical areas (e.g. auditory areas during motor activity, motor areas during sound listening) (A). Mean T-values of the ROI analysis for the "cismodally" stimulated cortical areas (e.g. motor areas during motor activity, auditory areas during sound listening) (B). Mean T-values of the VOLUNTARY mode are indicated by white bars, mean T-values of the INVOLUNTARY mode are indicated by black bars. Heschl's gyrus (HG), planum temporale (PT), planum polare (PP), primary motor cortex (M1), pre-supplementary motor area (pSMA), dorsal premotor cortex (PMd), ventral premotor cortex (PMv), nucleus lentiformis (NL). Errorbars indicate standard error. Locations of the "regions of interest" (ROI) are described in C. Displayed is only the right fraction of the x-axially symmetric ROIs. The ROIS are color coded as follows: PT (dark green), HG (magenta), PP (black), NL (cyan), M1 (red), PMd (yellow), PMv (blue), pSMA (light green).

Comparison between pianists and non-musicians

In order to study the effect of training on the proposed audio-motor network we compared activated areas in pianists and in non-musicians. By masking the areas that were concurrently present in auditory and motor conditions we focused on our proposed audio-motor network described in the *combined playing & listening vs. OFF* contrast. In addition, we focused on the modality that was not directly related to the task (e. g. the motor areas for auditory stimulation and the auditory areas for the motor activity respectively), because these areas rely on information from the other modality. Considering these constraints, we observed a general trend towards stronger activations in pianists. However, the condition when the subjects listened to the piece of music and were told to imagine the matching finger movements (and random finger movements for the non-musicians respectively) was the only one that showed significant differences ($p<0.05$, corrected). The left PMd and preSMA was more activated in pianists during listening to Mozart (Fig4, Tab. 2). In the INVOLUNTARY mode of the same condition, we observed an advantage for right PMd and SMA only on an uncorrected level ($p<0.005$). We observed no increased activity for non-musicians vs. musicians above the level of $p<0.005$, uncorrected in any of the conditions on the declared constraints. The maximal difference between the groups outside the described mask was observed in the superior parietal lobe in favor of the musicians.

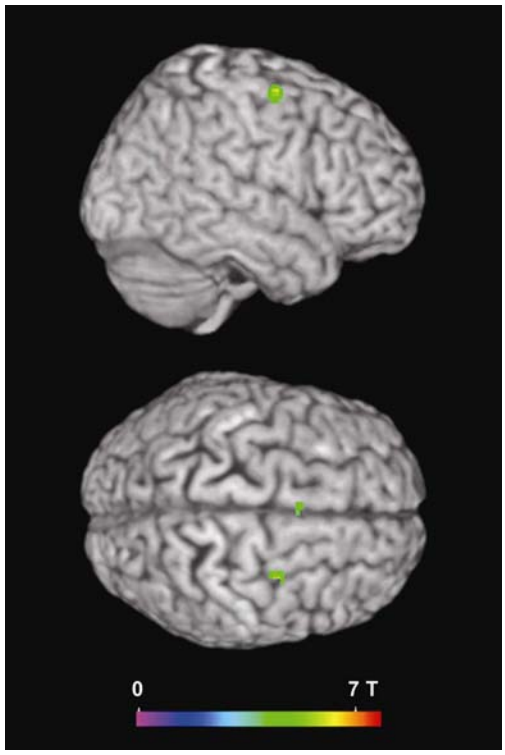


Fig. 4. Activations of the PIANISTS vs. NON-MUSICIANS contrast in the VOLUNTARY mode masked by the concurrently activated areas described in Fig. 2D ($p<0.05$, corrected for volume of mask).

6.5 Discussion

The presented data offer a set of presumably interconnected cortical areas engaged in both piano playing and piano listening. These cortical areas are ideal candidates for audio-motor integration and transformation circuits mediating information between the two modalities. Our discussion first considers the concurrently activated auditory and motor areas in pianists and non-musicians which appear in both VOLUNTARY and INVOLUNTARY modes. Then we discuss the role of voluntary involvement in audio-motor integration and the importance of cortical areas that are specifically activated in the VOLUNTARY mode. Finally, we focus on the differences in activation between pianists and non-musicians in the described network.

Transmodally activated auditory and motor areas

Our design and the presented data provide clear evidence that secondary auditory and motor areas namely the PMd, the PT and the PP exchange information during the multimodal task of piano performance. Assuming that auditory information in the pure auditory tasks enters the cortex first in primary auditory regions before being transferred to secondary auditory areas, additional activity in secondary motor areas implies direct or indirect information transfer from the auditory to the motor cortex. In the case of activation in the secondary auditory cortex during finger movement in silence the direction of information flow is less clear. Internally represented auditory information could influence motor execution via secondary and primary motor cortices or alternatively, the execution of movements could be echoed by information transfer to the secondary auditory cortex. In the case of piano performance both directions would make sense. The first may represent coordination of the finger movements with respect to an internal sound target representation (e.g. adjusting the finger force in respect to the internally represented sound level or even programming of sequences which will result in motor activity evoking the represented target sound). The second direction could provide motor information for the auditory cortex on the anticipated (internally generated) sound in order to compare it to external sound entering via the primary auditory cortex (this external sound is not present in our special case). As a matter of course, all these

considerations do not exclude multiple forward and backward flow of information between the auditory and motor systems as proposed by Burnod et al. (1999a). In any case, the observation that only PMd, PT and PP were significantly activated in the pure motor and pure auditory task during the INVOLUNTARY mode implies rather direct and automatically operating connections between these auditory and motor areas. Furthermore, at least two of the described areas, the PT and the PMd are repeatedly reported to be transmodally activated responding to the stimulation of various modalities (Buccino et al., 2001b; Foxe et al., 2002; Hasegawa et al., 2004; Haslinger et al., 2005; Petitto et al., 2000).

The proposed key-role of the PT and PMd in the audio-motor interaction pathway presents these structures as ideal candidates for audio-motor transformation circuits converting modality specific input from the primary cortices into a more general code that can be processed by other modalities. The existence of such audio-motor integration circuits were previously proposed by Doupe & Kuhl (1999) and Hickok & Poeppel (2000) for speech production. Hickok et al. (2003) and very recently Callan et al. (2006) observed activation in the premotor cortex and the planum temporale/superjor temporal parietal region for audio-motor language tasks. Based on language studies, Warren et al. (2005) proposed a general role for the PT/PMC in the much debated dorsal auditory pathway and presumed a similar involvement of these areas for instrumental music performance.

The third of the activated areas described above covers the anterior part of Brodmann Area (BA) 22 and the posterior part of BA38. We refer to this area as planum polare (PP) although the described activations are not confined to the superior temporal plane in all cases. The PP and BA38 respectively were described as mediating various auditory tasks (Jancke et al., 2001; Mazoyer et al., 1993; Meyer et al., 2000; Patterson et al., 2002; Warren et al., 2003) To the best of our knowledge, only two studies showed the PP activated in isolation (Brown et al., 2004; Zatorre and Belin, 2001). The first study is an investigation on the human song system which demonstrated isolated PP activation comparing cortical activations obtained during complex vocal melody reproduction with a simple vocal reproduction. From these findings the authors suggest that the PP is involved in higher-level music processing with special emphasis on the analysis if pitch and time structure. The second study compared a spectral with temporal processing task. The notion of PP as a higher-level cortical perception and production center involved in

music processing would ideally fit to the data of our experiment as the PP showed the most robust activation in all auditory and motor conditions of both the VOLUNTARY and INVOLUTARY modes. The involvement in perception or production of complex pitch and time structures with special emphasis on music is one of the few aspects that all these conditions have in common. Furthermore, this example provides additional evidence for the proposed similarity between human song production and instrumental music performance (Brown et al., 2004).

The role of the primary cortices

In contrast to electromagnetic studies (Haueisen and Knosche, 2001; Popescu et al., 2004) we observed no consistent activation in the primary motor cortex during sound listening. The same was true for the primary auditory cortex during finger movements (Tab. 1, 2, Fig. 2). Basically similar results were reported by two other studies which observed transmodal activation in the auditory cortex by visual stimulation. Hasegawa et al. (2004) and Haslinger et al. (2005) investigated indirect auditory and motor activation in musicians who observed silent video sequences of hands playing the piano. Both studies observed multiple activity clusters in secondary motor and auditory cortex (BA 6, 22, 42) bilaterally but no activation of the primary motor cortex. Only one of these studies observed unilateral activity in the primary auditory cortex. As mentioned above, we observed sporadically activated voxels in some of the subjects (musicians and non-musicians). However, the majority of the subjects investigated in our study even showed decreased activation in the primary motor cortex during auditory stimulation as reflected by negative mean t-values in the region of interest analysis (Shmuel et al., 2006) (Fig. 3A). Apart from potential methodological differences (BOLD- vs. electromagnetic signal) the most simple explanation for this disagreement refers to the limited spatial resolution and accuracy of electromagnetic recordings. The primary auditory and most notably the primary motor cortex are relatively narrow and encompassed by secondary areas. Thus, PET or fMRI are the methods of choice to distinguish between signals from primary and secondary areas.

Top down influence on the audio-motor network

Audio-motor activations in the INVOLUNTARY mode indicate the existence of an automated component in audio-motor integration. However, the pianists in our study reported that they were highly involved in both the auditory and the motor modality during piano performance. The results of the VOLUNTARY mode where the participants were requested to voluntarily imagine the absent modality (either the non-existing auditory stimuli during piano playing without hearing the tune and vice versa) support this notion. Compared to the involuntary involvement condition the transmodal activity is increased in all previously activated areas. Furthermore, there are several additional cortical areas concurrently activated during motor and auditory stimulation (Fig. 2D, Tab. 2). The fact that we observed activity in the PMv and the IFG (pars opercularis) exclusively in the VOLUNTARY mode is particularly interesting because these areas are among the key structures proposed for sensory-motor integration and transformation (Buccino et al., 2004; Rizzolatti et al., 1996; Romo et al., 2004; Schwartz et al., 2004). Our results, however, suggest that these structures are mainly involved in the voluntary, top down controlled component of this process. To the best of our knowledge this is the first study that shows activation of the PMv specifically to voluntary sensory-motor interaction but not to automatic sensory-motor interaction. Indirect support for this result is provided by a study of Cerasa et al. (2005). Using a finger tapping task they reported activity modulation in the PMv, the caudate nucleus of the basal ganglia and the IFG related to attention shifts while activation in the PMd, the SMA and the cerebellum was not significantly affected.

The activity of the cerebellum in the group-analysis is probably underestimated. Activity clusters in the cerebellum where consistently present in at least one hemisphere on a $p>0.05$ significance level, corrected for multiple comparisons on a single subject basis. The reason for the weak activation in the group-analysis is presumably an occasional cutoff of parts of the cerebellum due to a slight shift of the subjects head position during the measurement. Thus, the normalization procedure in this area was not optimal, leading to poor overlap between cerebellar activations of individuals.

Our observation that activity in the SMA, in contrast to the PMd, is more dependent on top down influences corresponds well with previous findings in humans and

animals that the SMA is more active during internally guided movements whereas the PMd is preferentially active guided by input from sensory modalities (Cunnington et al., 2002; Kurata and Wise, 1988; Mushiake et al., 1991; Thaler et al., 1988).
A further activated secondary area in our proposed audio-motor network is SII (parietal operculum). This cortical area is often considered as a potential link between the somatosensory cortex and other modalities, particularly the auditory system (Brett-Green et al., 2004; Jancke et al., 2006; Lutkenhoner et al., 2002; Menzel and Barth, 2005). Because each movement evokes somatosensory stimulation, in our case, SII may play a role in audio/motor-somatosensory transformation similar to the PT/PMC.
It might be argued that the presented transmodal activation during the VOLUNTARY mode is simply a reflection of imagery triggered by a cue from the other modality. This question was already addressed by Hickok et al. (2003). Therefore, we will not discuss this question in detail. However, we agree with the presented opinion that transmodally triggered imagery and multimodal integration are two aspects of the same process. In order to coordinate motor and auditory modalities in a feedback loop internal representations of the perceived sound and planed movement need to be compared to adjust sensory input and target output. Furthermore, a transformation of the representations from one modality into the other is required. It is likely that this kind of internal representation shares features and even circuits with ordinary imagery.

Musicians vs. non-musicians

The fact that we observed transmodal activity in all musicians and non-musicians during performance of piano related tasks suggests that this activity has a meaning beyond instrument playing per se since non-musicians never played an instrument and practiced scale playing just half an hour before the experiment. We rather suggest that this activity relates to a network that serves a wide range of audio-motor tasks. This opinion is supported by several studies which observed some of our proposed audio-motor areas active in audio-motor tasks not related to music (Buccino et al., 2001a; Jancke et al., 2000a; Kurata et al., 2000; Tettamanti et al., 2005). Furthermore, the most consistent activation in musicians and non-musicians was located bilaterally in PP, a structure that is not only tied to music performance

but also often reported in the context of language tasks. Language is probably the most commonly practiced audio-motor ability in humans. At this point it is also worth mentioning that most of the non-musicians reported uncertainties in the post-measurement-questionnaire in that they were not able to clearly indicate if they perceived a vocally sung melody instead of piano sound in the pure motor task. Furthermore, some pianists and non-musicians reported that they felt like subvocally singing along the piano sound. Therefore we can not completely exclude some subvocal audio-motor components responsible for the activations observed in our network.

Nevertheless, pianists compared to non-musicians showed increased transmodal activity in the motor part of the network for music listening. It is particularly interesting that apart from the PMd, we observed also increased transmodal activity in the preSMA, a structure which usually shows decreased activity in the comparison of pianist with non-musicans for piano motor execution or similar finger movements (Jancke et al., 2000b). This clearly suggests an increased motor representation for piano sound in pianists. Although the difference in secondary motor activity in pianists compared to non-musicians evoked by music is less distinctive when it occurs in an implicit manner (INVOLUNTARY mode) our data suggests that this difference is not completely driven by a top-down mode.

There is a slight imbalance in gender between the musicians and non-musicians. Possible gender effects in musical processing (Gaab et al., 2003) as well as in structural brain asymmetry (Luders et al., 2004) have been described. However, we do not think that any of the musician – non-musician differences could be explained by gender effects.

The fact that we observed no significant group difference for transmodal activity in the scale conditions and that non-musicians also showed considerable transmodal activation here indicates that audio motor coupling for simple tasks does not require years of training. The studies by Kurata et al. (2000) and very recently Lahav et al. (2005) which investigated short-term audio-motor conditioning/training support this view. It could also mean that the advantage of pianists in coupling music perception and finger movements is probably task specific and not necessarily transferable to audio motor tasks in general. However, our data do not allow to justify this claim and additional experiments testing different kinds of audio-motor tasks are required. Nevertheless, our data demonstrate that years of piano training lead to a modified

interaction of secondary motor areas with the auditory system. A notion which is greatly supported by the results of the recent work by Bangert et al. (2006).

Conclusion

We could clearly demonstrate that piano performance activates an audio-motor network of secondary and higher order structures which are tightly interconnected. This network shares many cortical structures also observed in visual-motor tasks. However, the observed network involves secondary temporal areas instead of parietal areas. Furthermore, we could show that audio-motor information flow consists of a voluntary and an involuntary component. These components make partially use of the same structures but voluntary involvement activates also additional cortical areas namely the ventral premotor areas and the pars opercularis of the IFG. Finally, we present evidence for long-term plasticity of the proposed audio-motor connections based on intensive piano training. Stronger activity is evoked in secondary motor cortices of pianists listening to music compared to non-musicians.

Acknowledgements

During preparation of this manuscript the authors were supported by Swiss National Science Foundation (SNF) SNF 46234101 (S.B. and M.M.) and by the SNF priority project NCCR/TP5 “Neural plasticity and repair” (S.K.).

6.6 References

Annett, M.,1970. A classification of hand preference by association analysis. Br. J. Psychol. 61, 303-321

Bangert, M. and Altenmuller, E.O.,2003. Mapping perception to action in piano practice: a longitudinal DC-EEG study. Bmc Neuroscience 4,

Bangert, M., Haeusler, U., Altenmuller, E.,2001. On practice: How the brain connects piano keys and piano sounds. Biological Foundations of Music 930, 425-428

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.J., Altenmuller, E.,2006. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. Neuroimage 30, 917-926

Brett-Green, B., Paulsen, M., Staba, R.J., Fifkova, E., Barth, D.S.,2004. Two distinct regions of secondary somatosensory cortex in the rat: Topographical organization and multisensory responses. Journal of Neurophysiology 91, 1327-1336

Brown, S., Martinez, M.J., Parsons, L.M.,2004. Passive music listening spontaneously engages limbic and paralimbic systems. Neuroreport 15, 2033-2037

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J.,2001a. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience 13, 400-404

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J.,2001b. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience 13, 400-404

Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.J., Rizzolatti, G.,2004. Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. Neuron 42, 323-334

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H.J., Jancke, L.,2005. Scanning silence: Mental imagery of complex sounds. Neuroimage 26, 1119-1127

Burnod, Y., Baraduc, P., Battaglia-Mayer, A., Guigon, E., Koechlin, E., Ferraina, S., Lacquaniti, F., Caminiti, R.,1999a. Parieto-frontal coding of reaching: an integrated framework. Experimental Brain Research 129, 325-346

Burnod, Y., Baraduc, P., Battaglia-Mayer, A., Guigon, E., Koechlin, E., Ferraina, S., Lacquaniti, F., Caminiti, R.,1999b. Parieto-frontal coding of reaching: an integrated framework. Experimental Brain Research 129, 325-346

Callan, D.E., Tsytsarev, V., Hanakawa, T., Callan, A.M., Katsuhara, M., Fukuyama, H., Turner, R.,2006. Song and speech: Brain regions involved with perception and covert production. Neuroimage

Caminiti, R., Johnson, P.B., Galli, C., Ferraina, S., Burnod, Y.,1991. Making Arm Movements Within Different Parts of Space - the Premotor and Motor Cortical Representation of A Coordinate System for Reaching to Visual Targets. Journal of Neuroscience 11, 1182-1197

Cerasa, A., Hagberg, G.E., Bianciardi, M., Sabatini, U.,2005. Visually cued motor synchronization: modulation of fMRI activation patterns by baseline condition. Neuroscience Letters 373, 32-37

Cunnington, R., Windischberger, C., Deecke, L., Moser, E.,2002. The preparation and execution of self-initiated and externally-triggered movement: A study of event-related fMRI. Neuroimage 15, 373-385

de Zwart, J.A., van Gelderen, P., Kellman, P., Duyn, J.H.,2002. Reduction of gradient acoustic noise in MRI using SENSE-EPI. Neuroimage 16, 1151-1155

Doupe, A.J. and Kuhl, P.K.,1999. Birdsong and human speech: Common themes and mechanisms. Annual Review of Neuroscience 22, 567-631

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E.,1995. Increased Cortical Representation of the Fingers of the Left Hand in String Players. Science 270, 305-307

Ellermann, J.M., Siegal, J.D., Strupp, J.P., Ebner, T.J., Ugurbil, K.,1998. Activation of visuomotor systems during visually guided movements: A functional MRI study. Journal of Magnetic Resonance 131, 272-285

Fink, G.R., Frackowiak, R.S., Pietrzyk, U., Passingham, R.E.,1997. Multiple nonprimary motor areas in the human cortex. J Neurophysiol. 77, 2164-2174

Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W., Murray, M.M.,2002. Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. Journal of Neurophysiology 88, 540-543

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.,1995. Statistical parametric maps in functional imaging: A General Linear Approach. Hum. Brain Mapp. 2, 189-210

Gaab, N., Keenan, J.P., Schlaug, G.,2003. The effects of gender on the neural substrates of pitch memory. J. Cogn Neurosci. 15, 810-820

Gaser, C. and Schlaug, G.,2003. Brain structures differ between musicians and non-musicians. Journal of Neuroscience 23, 9240-9245

Grefkes, C., Ritzl, A., Zilles, K., Fink, G.R.,2004. Human medial intraparietal cortex subserves visuomotor coordinate transformation. Neuroimage 23, 1494-1506

Halpern, A.R. and Zatorre, R.J.,1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. Cerebral Cortex 9, 697-704

Hasegawa, T., Matsuki, K.I., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., Sadato, N.,2004. Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. Cognitive Brain Research 20, 510-518

Haslinger, B., Erhard, P., Altenmuller, E., Schroeder, U., Boecker, H., Ceballos-Baumann, A.O.,2005. Transmodal sensorimotor networks during action observation in professional pianists. Journal of Cognitive Neuroscience 17, 282-293

Haueisen, J. and Knosche, T.R.,2001. Involuntary motor activity in pianists evoked by music perception. Journal of Cognitive Neuroscience 13, 786-792

Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T.,2003. Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. Journal of Cognitive Neuroscience 15, 673-682

Hickok, G. and Poeppel, D.,2000. Towards a functional neuroanatomy of speech perception. Trends in Cognitive Sciences 4, 131-138

Holm, S.,1979. A simple sequentially rejective multiple test procedure. Scan J Statist 6, 65-70

Hund-Georgiadis, M. and von Cramon, D.Y.,1999. Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. Experimental Brain Research 125, 417-425

Hutchinson, S., Lee, L.H.L., Gaab, N., Schlaug, G.,2003. Cerebellar volume of musicians. Cerebral Cortex 13, 943-949

Janata, P.,2001. Brain electrical activity evoked by mental formation of auditory expectations and images. Brain Topography 13, 169-193

Janata, P. and Grafton, S.T.,2003. Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. Nature Neuroscience 6, 682-687

Jancke, L.,1996. The hand performance test with a modified time limit instruction enables the examination of hand performance asymmetries in adults. Percept. Mot. Skills 82, 735-738

Jancke, L., Baumann, S., Koeneke, S., Meyer, M., Laeng, B., Peters, M., Lutz, K.,2006. Neural control of playing a reversed piano: empirical evidence for an unusual cortical organization of musical functions. Neuroreport 17, 447-451

Jancke, L., Buchanan, T.W., Lutz, K., Shah, N.J.,2001. Focused and nonfocused attention in verbal and emotional dichotic listening: An FMRI study. Brain and Language 78, 349-363

Jancke, L., Loose, R., Lutz, K., Specht, K., Shah, N.J.,2000a. Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. Cognitive Brain Research 10, 51-66

Jancke, L., Schlaug, G., Huang, Y.X., Steinmetz, H.,1994. Asymmetry of the Planum Parietale. Neuroreport 5, 1161-1163

Jancke, L., Shah, N.J., Peters, M.,2000b. Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. Cognitive Brain Research 10, 177-183

Kristeva, R., Chakarov, V., Schulte-Monting, E., Spreer, J.,2003. Activation of cortical areas in music execution and imagining: a high-resolution EEG study. Neuroimage 20, 1872-1883

Kurata, K., Tsuji, T., Naraki, S., Seino, M., Abe, Y.,2000. Activation of the dorsal premotor cortex and pre-supplementary motor area of humans during an auditory conditional motor task. Journal of Neurophysiology 84, 1667-1672

Kurata, K. and Wise, S.P.,1988. Premotor and Supplementary Motor Cortex in Rhesus-Monkeys - Neuronal-Activity During Externally-Instructed and Internally-Instructed Motor-Tasks. Experimental Brain Research 72, 237-248

Lahav, A., Boulanger, A., Schlaug, G., Saltzman, E.,2005. The power of listening: auditory-motor interactions in musical training. Ann. N. Y. Acad. Sci. 1060, 189-194

Lotze, M., Scheler, G., Tan, H.R.M., Braun, C., Birbaumer, N.,2003. The musician's brain: functional imaging of amateurs and professionals during performance and imagery. Neuroimage 20, 1817-1829

Luders, E., Narr, K.L., Thompson, P.M., Rex, D.E., Jancke, L., Steinmetz, H., Toga, A.W.,2004. Gender differences in cortical complexity. Nat. Neurosci. 7, 799-800

Lutkenhoner, B., Lammertmann, C., Simoes, C., Hari, R.,2002. Magnetoencephalographic correlates of audiotactile interaction. Neuroimage 15, 509-522

Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., Mehler, J.,1993. The Cortical Representation of Speech. Journal of Cognitive Neuroscience 5, 467-479

Meister, I.G., Krings, T., Foltys, H., Boroojerdi, B., Muller, M., Topper, R., Thron, A.,2004. Playing piano in the mind - an fMRI study on music imagery and performance in pianists. Cognitive Brain Research 19, 219-228

Menzel, R.R. and Barth, D.S.,2005. Multisensory and secondary somatosensory cortex in the rat. Cereb. Cortex 15, 1690-1696

Meyer, M., Friederici, A.D., von Cramon, D.Y.,2000. Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands. Cognitive Brain Research 9, 19-33

Munte, T.F., Altenmuller, E., Jancke, L.,2002. The musician's brain as a model of neuroplasticity. Nature Reviews Neuroscience 3, 473-478

Mushiake, H., Inase, M., Tanji, J.,1991. Neuronal-Activity in the Primate Premotor, Supplementary, and Precentral Motor Cortex During Visually Guided and Internally Determined Sequential Movements. Journal of Neurophysiology 66, 705-718

Pantev, C., Roberts, L.E., Schulz, M., Engelien, A., Ross, B.,2001. Timbre-specific enhancement of auditory cortical representations in musicians. Neuroreport 12, 169-174

Parsons, L.M., Sergent, J., Hodges, D.A., Fox, P.T.,2005. The brain basis of piano performance. Neuropsychologia 43, 199-215

Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., Griffiths, T.D.,2002. The processing of temporal pitch and melody information in auditory cortex. Neuron 36, 767-776

Petitto, L.A., Zatorre, R.J., Gauna, K., Nikelski, E.J., Dostie, D., Evans, A.C.,2000. Speech-like cerebral activity in profoundly deaf people processing signed languages: Implications for the neural basis of human language. Proceedings of the National Academy of Sciences of the United States of America 97, 13961-13966

Popescu, M., Otsuka, A., Ioannides, A.A.,2004. Dynamics of brain activity in motor and frontal cortical areas during music listening: a magnetoencephalographic study. Neuroimage 21, 1622-1638

Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P.,1999. SENSE: Sensitivity encoding for fast MRI. Magnetic Resonance in Medicine 42, 952-962

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.,1996. Premotor cortex and the recognition of motor actions. Cognitive Brain Research 3, 131-141

Roland, P.E. and Zilles, K.,1996. Functions and structures of the motor cortices in humans. Curr. Opin. Neurobiol. 6, 773-781

Romo, R., Hernandez, A., Zainos, A.,2004. Neuronal correlates of a perceptual decision in ventral premotor cortex. Neuron 41, 165-173

Schlaug, G.,2001. The brain of musicians: A model for functional and structural adaptation. Biological Foundations of Music 930, 281-299

Schlaug, G., Jancke, L., Huang, Y.X., Staiger, J.F., Steinmetz, H.,1995. Increased Corpus-Callosum Size in Musicians. Neuropsychologia 33, 1047-&

Schmahmann, J.D., Doyon, J., Evans, A.C., Petrides, M., and Toga, A.W., 2000. MRI Atlas of the Human Cerebellum. In: Academic Press, San Diego, pp

Schwartz, A.B., Moran, D.W., Reina, G.A.,2004. Differential representation of perception and action in the frontal cortex. Science 303, 380-383

Shmuel, A., Augath, M., Oeltermann, A., Logothetis, N.K.,2006. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. Nature Neuroscience 9, 569-577

Steingruber, H.J.,1971. [Measuring of laterality]. Z. Exp. Angew. Psychol. 18, 337-357

Steinmetz, H., Rademacher, J., Jancke, L., Huang, Y.X., Thron, A., Zilles, K.,1990. Total Surface of Temporoparietal Intrasylvian Cortex - Diverging Left Right Asymmetries. Brain and Language 39, 357-372

Steinmetz, H., Staiger, J.F., Schlaug, G., Huang, Y.X., Jancke, L.,1996. Inverse relationship between brain size and callosal connectivity. Naturwissenschaften 83, 221-221

Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S.F., Perani, D.,2005. Listening to action-related sentences activates fronto-parietal motor circuits. Journal of Cognitive Neuroscience 17, 273-281

Thaler, D.E., Rolls, E.T., Passingham, R.E.,1988. Neuronal-Activity of the Supplementary Motor Area (Sma) During Internally and Externally Triggered Wrist Movements. Neuroscience Letters 93, 264-269

Warren, J.D., Uppenkamp, S., Patterson, R.D., Griffiths, T.D.,2003. Separating pitch chroma and pitch height in the human brain. Proceedings of the National Academy of Sciences of the United States of America 100, 10038-10042

Warren, J.E., Wise, R.J., Warren, J.D.,2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci. 28, 636-643

Zatorre, R.J. and Belin, P.,2001. Spectral and temporal processing in human auditory cortex. Cerebral Cortex 11, 946-953

Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., Evans, A.C.,1996. Hearing in the mind's ear: A PET investigation of musical imagery and perception. Journal of Cognitive Neuroscience 8, 29-46

7 Fourth Study

Audio-Visual Intertwining in the Human Brain during Crossmodal Associative Learning

MARTIN MEYER, SIMON BAUMANN, SARAH MARCHINA & LUTZ JÄNCKE

Div. Neuropsychology, Institute for Psychology, University of Zurich

Submitted to *Neuroscience Research*

7.1 Abstract

Fifteen healthy participants partake in a "silent" sparse temporal event-related fMRI study to explore neural substrates of rapid cross-modal associative learning. Therefore the participants were presented with paired (visual and auditory) and unpaired (only visual following the paired presentation) stimulation. Another unpaired visual stimulation preceding the paired stimulation should serve as control condition. During both unpaired visual presentations (preceding and following the paired presentation) we observed significant brain responses beyond primary visual cortex in the bilateral auditory association cortex (planum temporale), in the right inferior parietal lobe, and in the bilateral insulae. This result implies that the consistent occurrence of visual cues and subsequent acoustic scanner noise apparently triggers activation in associative cortices. More generally, our study indicates the existence of tightly intertwined visual-auditory networks since anticipating scanner noise apparently provokes hemodynamic responses in associative auditory cortices reecting triggered imagery but not sensation of an auditory event. Evidently, our

findings show that in humans, even simple pairing of auditory and visual stimuli induces bottom-up associative learning.

7.2 Introduction

Associative learning and adaptation could be regarded as one of the most fundamental behavioural functions in both humans and animals. The processing of external and internal sensations is an important prerequisite for learning. Feeling the painful heat while viewing and touching a hot plate helps an organism learn about a potential danger. In this example, the combination of multiple visual and somatosensory sensations enables an organism to establish an association between an object, the hot plate, and a painful action and thus helps learn an appropriate behaviour. But even paired sensations that are not this unpleasant clearly demonstrate that multisensory learning may establish a strong relationship between two events. Seeing lightning makes an individual anticipate hearing thunder. Thus, inputs from the different sensory modalities are combined to form a single integrated experience of the world (Ghazanfar & Schroeder, 2006). Multisensory sensations and integrations are enormously important and advantageous tools in an organism's repertoire to effectively learn how to act properly and how to avoid deleterious experiences. However, to date little is known about the neural underpinnings of multisensory learning in the human brain. A modicum of studies has demonstrated the existence of functional amalgamations and structural connections across modalities which are supposed to constitute basic mechanisms of learning (Calvert et al., 2000; Calvert, 2001). As highlighted by Molholm & Foxe (2005, p. 123) it is still an open question how inputs from various sensory systems precisely interact across the ascending levels of the cortical and (and perhaps even subcortical) processing hierarchies. In respect to cortico-subcortical networks it has been observed that visual cues influence auditory responses in the rat thalamus (Komura et al., 2005). With regard to cortical networks there is abundant evidence indicating that primary and associative sensory regions preferentially bind to together to enable multisensory learning. Using functional magnetic resonance imaging (fMRI) Foxe et al. (2002) demonstrated that auditory and somatosensory inputs converge in a subregion of

human auditory cortex along the superior temporal gyrus (STG). Schroeder and colleagues underscore this observation by providing anatomical evidence of visual and somatosensory input into putatively unisensory regions at the secondary/tertiary levels of the auditory hierarchy (Schroeder & Foxe, 2002; Schroeder et al., 2003). At the neurofunctional level recent electrophysiological studies using intracranial recordings from humans and animals have shown direct visual and somatosensory input to the caudomedial belt area of auditory association cortex (Schroeder et al., 2001; Schroeder & Foxe, 2002, 2005). Multisensory processing has also been subject to imaging studies on visual memory retrieval (Nyberg et al., 2000). In this study the authors observed that visual retrieval of auditorily presented words activates the auditory cortex. Murray and colleagues demonstrated that picture presentation paired with sounds results in improved memory performance (Murray et al., 2004). These multisensory memory representations are established even after single-trial exposure and are later accessible to facilitate memory which implies an extremely fast and robust establishment of multisensory representations (Murray et al., 2005). Based on this compelling evidence it is indisputable that the auditory cortex receives converging input from the visual system, and thus may be recruited in visual information processing (Schroeder & Foxe, 2005). Recent brain imaging studies testing human participants deliver further buttressing for the view that human auditory cortex may support processing of visual stimuli. In the context of language comprehension imaging studies report auditory cortex engagement during visual speech perception in the absence of spoken language, in particular when participants viewed articulatory gestures (Pekkola et al., 2005) or processed visual cues during silent lip-reading (Bernstein et al., 2002; Calvert et al., 1997; MacSweeney et al., 2000). Recent investigations on auditory imagery evoked by visual cues have also shown that mental imagery of complex sounds and short syllables brings on activation increases in primary (Yoo et al., 2001) and secondary auditory fields (Bunzeck et al., 2005; Jancke & Shah, 2004; Yoo et al., 2001). Results of another fMRI-study indicated that recalling learned sounds yields enhanced activation in human auditory association cortex (Wheeler et al., 2000).

All these aforementioned studies have in common that they encourage participants to embark on a controlled top-down strategy. However, it has also been shown that automatic bottom-up processing also triggers audio-visual intertwining. One seminal fMRI-study uncovered responses in the visual cortex to presentation of sounds in

isolation following a learning period in which a visual stimulus was consistently paired with an audible tone (McIntosh et al., 1998). Here, multisensory interactions were characterized in human subjects as they learned that an auditory stimulus signals a visual event. This study demonstrates to what extent parts of the nervous system operate as a cohesive system in learning about and responding to the environment. Based on this accrual of evidence that coherently suggests the existence of multisensory neuroplasticity in hetero-modal perception we surmise that information from auditory and visual networks are blended which forms the basis for rapid cross-modal perceptual learning, i.e. classical conditioning.

The present study

Thus, we set out a study involving human participants which is similar to the aforementioned experiment by McIntosh and colleagues, but addresses the question whether visual stimuli may induce activity in auditory association cortices after they had been presented in combination with sounds. We used a conditioning paradigm which should be suitable to investigate the simplest form of associative learning. This form of learning establishes a short-term relationship between two events even when these events are affectively neutral and do not have any relevance for the organism that undergoes conditioning. This form of learning occurs in any case when a previously neutral stimulus (CS) is temporally paired with another unconditioned stimulus (UCS) that evokes a physiological and/or behavioural response (UCR). After a phase of consistent temporally paired stimulation (short-delay conditioning) it suffices to present the formerly neutral stimulus (now CS+) to observe the response initially elicited by the UCS. The instantaneous forming of an association even functions when the CS and UCS do not have any natural linkage or when CS is presented outside of awareness (Büchel & Dolan, 2000).

Even though animal studies of classical conditioning have provided considerable information on the neural underpinnings of associative learning (Bao et al., 2004; Moucha et al., 2005), only a little research has been undertaken so far to reveal the neural ensembles involved in bottom-up multisensory associative learning in humans. Without any exception these studies examined classical defense conditioning (CDC)

which means that UCS is an aversive stimulus (electric shock, aversive sound etc.) (Büchel et al., 1998; Hugdahl et al., 1995; Molchan et al., 1994; Morris et al., 1998). The results of Hugdahl and colleagues is of particular note, they did not find any involvement of auditory cortex in response to conditioned aversive sounds but pronounced right inferior frontal activity which they consider to constitute an 'expectancy circuit'.

Based on previous findings from neuroimaging studies on bottom-up multisensory processing (McIntosh et al., 1998) we proposed an involvement of auditory association areas triggered by the presentation of visual stimuli in isolation which have precedingly been paired with auditory stimuli. Secondly, based on the results of the aforementioned PET-studies on classical conditioning we conjectured that inferior frontal regions are supposed to support associative learning as the inferior frontal cortex has been described as an additional supramodal resource which helps establish functional relationships in cross-modal conditioning (Hugdahl et al., 1995). Unlike previous imaging studies which made use of aversive auditory and tactile stimuli (Büchel et al., 1998; Hugdahl et al., 1995; Morris et al., 1998; Wik et al., 1997), we applied non-aversive simple sensory stimuli to avoid conobserveding with emotional processes.

To avoid a perceptual and physiological masking of auditory processing induced by scanner noise we applied a "silent" fMRI protocol (clustered-sparse temporal acquisition scheme, CTA). These schemes have been adopted to investigate the neural correlates of human audition using blood-oxygen-level dependent (BOLD) based functional magnetic resonance imaging (fMRI) devoid of ambient conobserveding acoustic scanner noise. Sparse temporal sampling has previously been extended to clustered-sparse temporal acquisition designs which record several subsequent BOLD contrast images in rapid succession, in order to enhance temporal sampling efficiency (for details see Materials and Methods section).

7.3 Materials and Methods

Participants. Sixteen healthy volunteers (8 males, 8 females, age range 24-40, mean 27.8 yrs.), all strongly right-handed according to a standard questionnaire (Annett, 1970; Jancke, 1996), partook in the study. Volunteers were not familiarized with the stimuli or procedure prior to scanning. They had no neurological or psychiatric history, nor did they have any visual or hearing disorder. Written informed consent was obtained prior to the examination. The study was in accordance with Zurich Medical Faculty Ethical guidelines. Due to motion artefacts one participant had to be excluded from analysis.

Stimuli and control. The study comprised one visual (CS) and one auditory (UCS) stimulus. The visual stimuli were either presented in isolation or paired with the auditory stimuli. We used a total screen red flash which lighted up for 100 ms followed by total dark screen (100 ms) which was again replaced by a red flash (100 ms). A telephone ringing[20] served as auditory stimulus. The sound signal was digitized at a 16 bit/44.1 kHz sampling rate and shortened to 2.6 s using the Magix Deluxe software (www.magix.com). Stimuli were controlled using Presentation software (Version 0.70, www.neurobs.com). Stimulus presentation was synchronized by a 5 V TTL trigger pulse with the data acquisition. We used standard Phillips headphones for binaural stimulus delivery. Zero trials during which neither auditory nor visual stimuli were presented served as silent control for data analysis. During null events participants viewed a black screen throughout the entire trial.

Procedure and task. Prior to scanning participants were informed about the experimental procedure but not about the scientific background of the study. Volunteers' task was to attend to the stimuli and to press a button alternately with the right and left hand's index finger after each trial signaled by the offset of scanner noise. As associative learning is supposed to occur automatically we involved our participants in the simple task, specifically not directing the subjects' attention to the

[20] MP3 download, www.jamba.de/dew/home.do

stimuli. It has been shown that explicitly directing a person's attention toward an auditory stimulus yields increased activation in auditory fields (Jancke et al., 1999). The task was designed to keep them generally attentive. Participants were comfortably placed supine in the scanner and underwent four experimental blocks separated by three short breaks during which subjects kept lying in the scanner. Each block corresponded to one particular experimental phase that we introduce in turn.

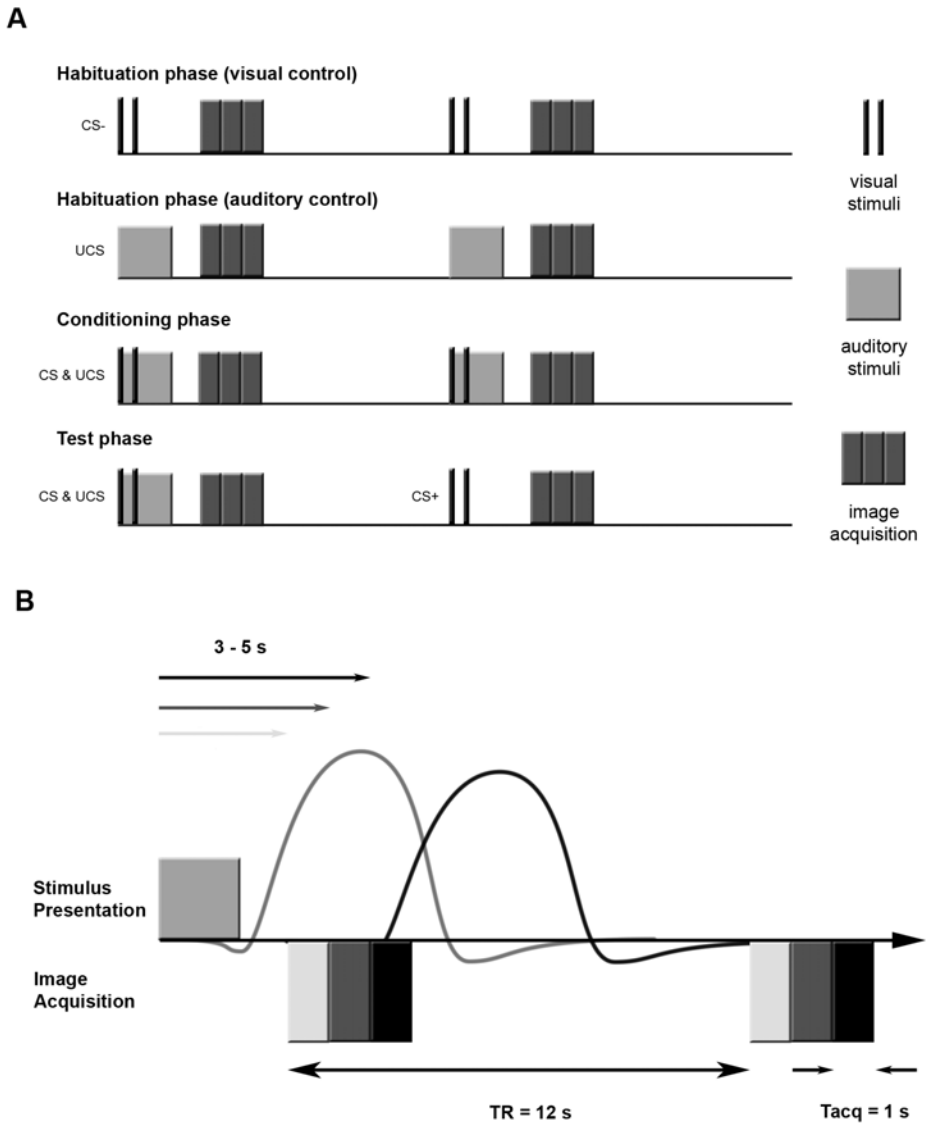


Fig. 1. (A) Schematic illustration of the four experimental phases. For details see the Methods section. (B) Schematic illustration of clustered temporal acquisition (CTA) as implemented in the present study. In each single trial the hemodynamic response is obtained from three consecutive volume scans (TR 1 s). The interval between onset of one volume triplet and onset of the subsequent triplet is 12 s. The interval between onset of stimulus presentation and data collection varies between 3 and 5 sec which allows enhanced sampling of data points relative to single volume acquisition. This approach precludes confounding of hemodynamic responses to stimuli and scanner noise.

The first habituation phase was supposed to serve as a visual control condition as participants only viewed visual stimuli in isolation (CS, n=32) and zero trials (n=16). The second habituation phase was meant to serve as an auditory control condition since volunteers only heard auditory stimuli in isolation (UCS, n=32) and zero trials (n=16). Associative learning was presumed to occur during the third phase whilst we consistently presented paired visual and auditory stimuli (CS and UCS, n=32) and zero trials (n=16). The fourth phase was targeted at testing to what extent an associative relationship between CS and UCS had been established. Participants were either presented with paired visual and auditory stimuli (CS and UCS, n=32) as in the preceding phase, visual stimuli in isolation (CS+, n=32), or zero trials (n=32). In other words, we applied a 5:10 reinforcement plan to partly maintain conditioning and to preclude fast extinction of the established association. While the duration of first, second, and third block was 12 minutes each, the scanning of the last phase took 24 minutes summing up to a total of 60 minutes scanning time for the functional part of the experiment. All participants experienced the same order of experimental phases. Generally, the sequence of visual, auditory, and empty presentation was pseudo-randomized within each block to preclude predictability.

Design. We implemented a clustered sparse temporal acquisition (CTA) technique that combines the principle design of a sparse temporal acquisition (STA) with the clustered acquisition of three consecutive volume scans per trial (Schmidt et al., submitted). Both CTA and STA method have been developed to overcome the problem that continuous acoustic noise resulting from switching of the magnet's gradient coils produces activation in the auditory cortex and hence seriously interferes with responses to acoustic stimuli in an auditory paradigm. Basically, CTA and STA are advantageous compared with traditional continuous scanning as they allow the presentation of auditory stimuli in a silent period prior to the onset of the acoustic noise emitted by switching gradients.

Thus, the design of the present study allows the analysis of hemodynamic responses which are not contaminated by brain responses to scanner noise. Our CTA scheme comprised the acquisition of three consecutive volume scans. A long inter-scan interval (repetition time 15 s) then allows both the functional response to the auditory stimulus and the response evoked by the scanner noise to decay prior to the next trial. This approach is capable of clearly separating the task-induced functional response from the scanner-noise induced functional response.

Data acquisition. Data were collected using a Philips Intera 3 T whole body MR unit (Philips Medical Systems, Best, The Netherlands) equipped with an eight-channel Philips SENSE head coil. Functional time series were obtained from 14 transverse slices covering auditory cortex with a spatial resolution of 2.7 x 2.7 x 4 mm3 using a Sensitivity Encoded (SENSE; Pruessmann et al., 1999) single-shot gradient-echo planar sequence (acquisition matrix 80 x 80, SENSE acceleration factor R 2.0, FOV 220 mm, TR 1000 ms, TE 35 ms and ip angle 90°). We collected three subsequent volumes during each trial which covered the functional response to the stimulus at the plateau of the hemodynamic response with a delay of 3 s, 4 s and 5 s after stimulus onset (Bandettini et al., 1998; Belin et al., 1999; Edmister et al., 1999; Hall et al., 1999) (cf. Figure 1B). Additionally, we obtained one echo planar image that covered the whole brain with 38 transverse slices (TR 4000 ms) but used otherwise the identical scan parameters as for the functional time series. This whole-head EPI volume was used to assist the spatial normalization of the functional time series (c.f. Data Analysis). Furthermore, we collected a standard 3D T1 weighted scan for anatomical reference with 1 x 1 x 0.8 mm3 spatial resolution (acquisition matrix 224 x 224, TE 2.30 ms, TR 20 ms, ip angle 20°).

Data analysis. To account for different T1 saturation effects in subsequent volumes, we subjected the three volume scans collected during each cluster to three separate time series during data analysis. Consequently, even if the longitudinal magnetization is partially saturated in the later volumes of the cluster, we contrasted the activation and baseline signals only from the "same-volume" time series. Furthermore, each of these three time-series corresponds to the hemodynamic response sampled at a distinct temporal window, i.e. 3 s, 4 s and 5 s after stimulus onset.

Post-processing and statistical analyses were carried out using MATLAB 6.5 (Mathworks Inc., Natiek, MA, USA) and the SPM99 software package (http://www.fil.ion.ucl.ac.uk/spm/). All volumes were realigned to the first volume, corrected for motion artefacts, mean-adjusted by proportional scaling, normalized into standard stereotactic space (template provided by the Montreal Neurological Institute). In order to optimize normalization we coregistered the functional time-series with the whole-head EPI-T1 images. For spatial smoothing we applied an isotropic Gaussian kernel (8 mm full-width-at-half-maximum). Low-frequency drifts were removed using a temporal high-pass filter (cut-off of 100 s).

Statistical analysis was based on the General Linear Model (Friston et al., 1995). Single trials were treated as epochs and modelled by means of a box car function. We calculated contrast images from each of the three volumes. The resulting set of voxel values for each contrast constitutes a statistical parametric map of the T-statistic [SPM(T)]. In order to explore the group-level activation across the 15 participants we used a random effects model (second level analysis on contrast images obtained from individuals). This model estimates the error variance for each condition across individual subjects rather than across all scans and this provides stronger generalization of the statistical population. For report and discussion of results only significant clusters of activation were considered (uncorrected α-level 0.001, $k \geq 10$).

We also performed a post hoc 'region of interest' (ROI) analysis which enabled us to test whether OLD responses obtained from distinct sites of the fronto-temporo-parietal cortex may vary as function of condition. For four conditions (visual habituation, auditory habituation, paired audiovisual presentation, and extinction) we collected BOLD signals recorded during the second out of three volumes from four bilateral ROIs placed in the insula, the mid portion of the STG, he planum temporale, and in the supramarginal gyrus (SMG) overarching the planum parietale (PPa) from all participants. Spherical ROIs (radius 3 mm) were defined as this approach guarantees homogeneity of variance due to the equal size of ROIs (Bosch, 2000). We defined coordinates of averaged local response maxima as centre voxels of ROIs (cf. Tables 1 - 4): LH STR/HG (-40, -28, 7), RH STR/HG (49, -18, 5), SMG/PPa (-60, -21, 27), RH SMG/PPa (63, -36, 27), LH PT (-62, -38, 18), RH PT (62, -38, 18). t-values were averaged within each distinct spherical ROI, across participants and hemispheres and subjected to systematic paired comparisons.

7.4 Results

Figure 3 displays the main fMRI results. Due to unsteady magnetization associated with the clustered temporal acquisition we only report activity collected with the second out of the three clustered trials. All main contrasts we report here are derived from the comparison between the experimental phases vs. silent control (zero trials) to elucidate the entire network of involved rain areas.

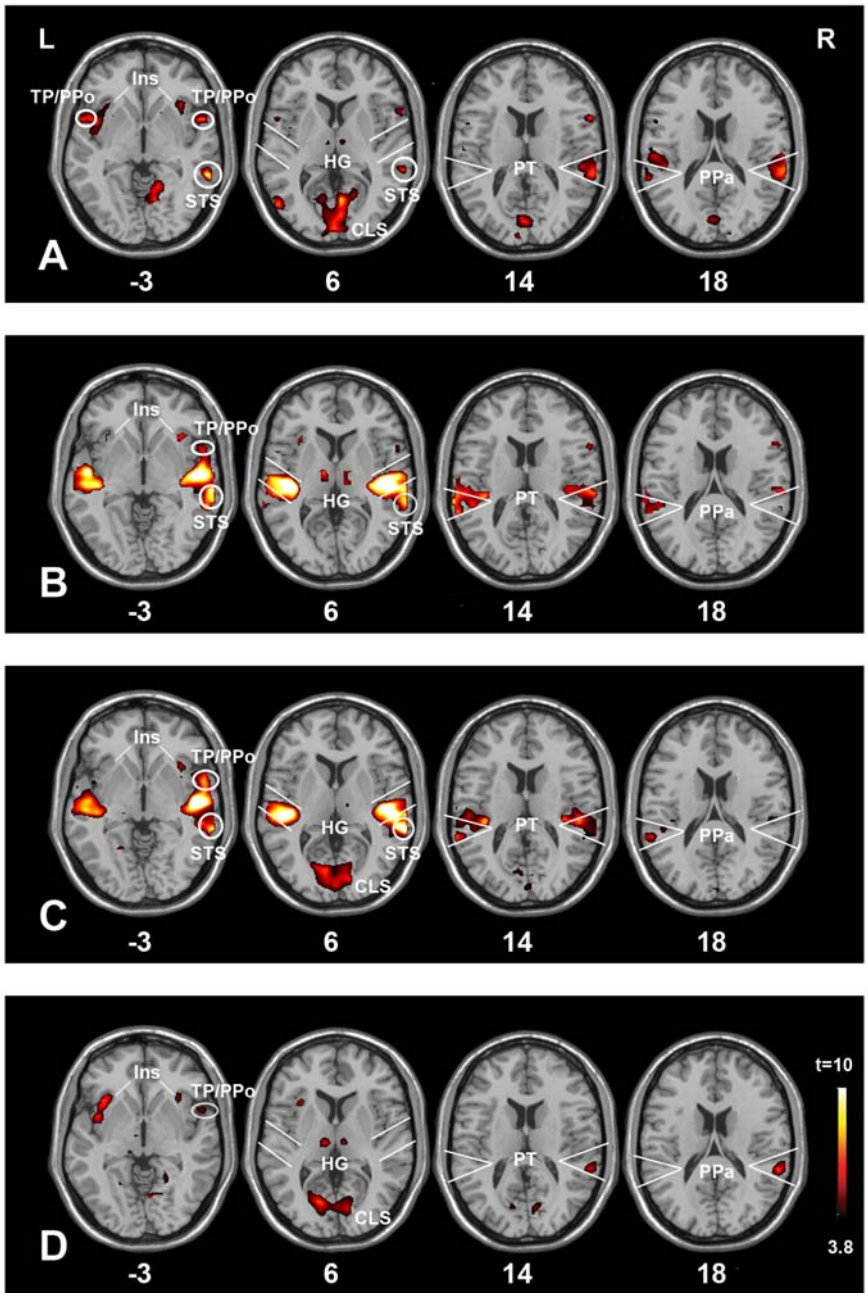


Fig. 3. Functional brain responses collected during succeeding experimental phases are depicted. The brain scans show consistently stronger functional activation for stimuli conditions relative to silent control obtained from the second out three volumes. All functional contrasts are thresholded at T = 3:79, p < 0.001 (uncorrected t-level, k > 10) and superimposed on transverse and sagittal slices of the MNI-T1-weighted standard brain. Tables 1-4 list peak activations (T-values) of distinct activation clusters and anatomical areas. [A] Visual habituation, [B] Auditory habituation, [C] Conditioning phase, [D] Test phase (extinction).

Table 1: **Vision vs. Silent control.** Local response maxima of significant clusters (random-effects analysis, $p \leq 0.001$ uncorrected for multiple comparisons; $k \geq 10$). Localization of clusters correspond to position of local maximal activations in normalized space of the MNI standard brain (www.mni.mcgill.ca) indicated by the Z value in a particular anatomical structure. Distances are relative to the intercommissural (AC-PC) line in the horizontal (x), anterior-posterior (y) and vertical (z) directions. We used the 'Automatic anatomical labeling' tool (Tzourio-Mazoyer et al., 2002) available for implementation in SPM99 (http://www.cyceron.fr/freeware/). Anatomical abbreviations are stipulated as follows: IFG(op) = inferior frontal gyrus (opercular part), IFG(tr) = inferior frontal gyrus (triangular part), HG = Heschl's gyrus, STG = superior temporal gyrus, STS = superior temporal sulcus, PT = planum temporale, PPa = planum parietale, MTG = middle temporal gyrus, SMG = supramarginal gyrus, TP = temporal pole, MOG = middle occipital gyrus, LG = lingual gyrus, CLS = calcarine sulcus, Cun = Cuneus, PCun = Precuneus, Ins = Insula, Tha = Thalamus, IC = inferior colliculus.

Vision > Silent Control	Left hemisphere					Right hemisphere				
Location	k	T value	x	y	z	k	T value	x	y	z
CLS/LG/Cun	511	4.93	0	-93	3	-	-	-	-	-
MTG/MOG	36	4.96	-54	-69	3	-	-	-	-	-
SMG/PPa	162	4.47	-60	-21	27	a	4.69	63	-36	27
STG/PT[b]	-	-	-	-	-	a	4.46	63	-36	15
STS[c]	-	-	-	-	-	263	5.17	60	-42	-3
Ins/TP	197	4.46	-42	15	-6	184	4.67	57	9	-6
Ins	-	-	-	-	-	d	4.23	40	24	-12
IFG(op)	-	-	-	-	-	28	4.52	60	15	12
Put	24	4.13	-24	12	-9	-	-	-	-	-
Tha	10	3.90	-3	-12	9	-	-	-	-	-

[a] This cluster cannot be distinctively separated by the cluster in the right STG. Thus this cluster has a local maximum, but no distinct extension.

[b] According to the probability atlas by Westbury et al. (1999) the local maximum of this cluster is situated in the PT with a 5-25% probability.

[c] According to the probability atlas by Westbury et al. (1999) the local maximum of this cluster is situated outside the PT.

[d] This cluster cannot be distinctively separated by the cluster in the adjacent TP. Thus this cluster has a local maximum, but no distinct extension.

Table 2: **Audition vs. Silent control.** Local response maxima of significant clusters (random-effects analysis, $p \leq 0.001$ uncorrected for multiple comparisons; $k \geq 10$). For explanations, see Table 1.

Audition > Silent Control	Left hemisphere					Right hemisphere				
Location	k	T value	x	y	z	k	T value	x	y	z
STG[a]	789	6.62	-36	-30	9	1003	6.73	48	-18	6
Ins	24	3.92	-30	24	3	24	4.08	36	21	-3
Tha	26	4.40	-9	-9	6	33	3.92	9	-18	3

[a] The cluster in the right hemisphere also encompassed the Rolandic operculum. According to the probability atlas by Rademacher et al. (2001) the local maximum of this cluster is situated in the HG with 40% probability.

Table 3: **Paired vision (CS) and audition (UCS) vs. Silent control .** Local response maxima of significant clusters (random-effects analysis, $p \leq 0.001$ uncorrected for multiple comparisons; $k \geq 10$). For explanations, see Table 1.

CS and UCS > Silent Control	Left hemisphere					Right hemisphere				
Location	k	T value	x	y	z	k	T value	x	y	z
STG/Ins	600	5.68	-45	-27	6	857	6.54	51	-18	3
CLS[a]	274	4.42	-15	-75	9	-	-	-	-	-
LG	14	3.72	-21	-57	0	-	-	-	-	-
Cun[a]	-	-	-	-	-	90	4.37	9	-75	39
IC	-	-	-	-	-	15	4.14	9	-24	-9

[a] Activation occurs in both left and right hemisphere with only the maximum T-value peaking in the indicated hemisphere.

Table 4: **Vision (CS+) vs. Silent control.** Local response maxima of significant clusters (random-effects analysis, $p \leq 0.001$ uncorrected for multiple comparisons; $k \geq 10$). For explanations, see Table 1.

CS+ > Silent Control	Left hemisphere					Right hemisphere				
Location	k	T value	x	y	z	k	T value	x	y	z
STG/PT[a]	-	-	-	-	-	84	4.45	60	-39	21
IFG(tri)	-	-	-	-	-	27	4.09	48	12	27
Ins	98	5.26	-30	21	3	14	3.74	33	21	0
Ins/TP	-	-	-	-	-	31	4.08	51	9	-9
CLS[b]	214	4.30	-15	-72	6	-	-	-	-	-
LG	-	-	-	-	-	16	4.37	21	-51	-6
PCun/Cun	-	-	-	-	-	20	3.69	9	-75	39
Tha	24	3.90	-9	-18	6	-	-	-	-	-
IC	16	4.63	-9	-27	-9	-	-	-	-	-

[a] According to the probability atlas by Westbury et al. (1999) the local maximum of this cluster is situated in the PT with 26-65% probability.
[b] Activation occurs in both left and right hemisphere with only the maximum T-value peaking in the indicated hemisphere.

First habituation phase (visual control) - Figure 3A and Table 1 illustrate that perception of visual stimuli was associated with activity in several cortical areas normally attributed to visual processing, namely the calcarine sulcus (CLS), the cuneus (Cun), and the left temporooccipital lobe (V4). We also noticed an engagement of the bilateral supramarginal gyrus (SMG) which overarches the planum parietale (PPa), the left posterior superior temporal gyrus including the planum temporale (PT) and the left posterior portion of the superior temporal sulcus (STS). Furthermore, we observed bilateral activity in supramodal regions involving the opercular part of the inferior frontal gyrus (IFG(op)), the anterior insula, the temporal pole, and the adjacent Rolandic operculum. Finally, the analysis revealed a recruitment of subcortical thalamic and basal ganglia regions (Putamen).

Second habituation phase (auditory control) - Figure 3B and Table 2 show that hearing simple telephone ringing yields considerable signal increase in primary and associative auditory fields stretching along the entire ventral bank of the peri-sylvian region including the supratemporal plane as well as the lateral STG. Furthermore, we uncovered small patches of activity in left and right anterior insulae as well as in thalamic regions.

Conditioning phase (paired vision (CS) and audition (UCS)) - Figure 3C and Table 3 depict brain responses while participants were presented with paired visual and auditory stimuli. The analysis identified significantly stimulated regions in the superior temporal region and in the anterior insulae bilaterally, in the calcarine sulcus, in the cuneus and in the inferior colliculus.

Maintenance (paired vision (CS-) and audition (UCS)) - We do not explicitly report activation in response of this region as maintenance turned out to involve the same regions as the preceding conditioning phase.

Test phase or extinction (only vision (CS+) - Figure 3D and Table 4 visualizes which areas were active while participants only viewed stimuli after conditioning had occurred. Besides responses in the visual cortex (CLS), in the right lingual gyrus, and in the right cuneus/precuneus we also identified right lateralized activity in auditory association regions of the posterior STG partly encroaching onto the PT. In the right hemisphere we also observed an activation cluster which covered the IFG(op) and the anterior temporal plane. Furthermore we observed bilateral responses in the anterior insula. Finally, this condition also brought on increased blood supply in subcortical regions, namely the thalamus and the inferior colliculus.

Post-hoc analysis.

Figure 4 and Table 5 show that mean t-values in temporal and parietal sites differ as a function of the experimental phase.

We subjected t-values to systematic comparisons (paired-samples t-test) within each ROI separately. These analyses attest the results of the statistical parametric maps that the involvement of distinct ROIs differs between regions as a function of condition. First, this analysis can also be considered another indication that the core auditory region is only engaged when participants listened to auditory stimuli. By contrast the statistical comparisons evidence that the PPa is most strongly involved during the visual control while activity in the PT is markedly larger during the last (visual) run.

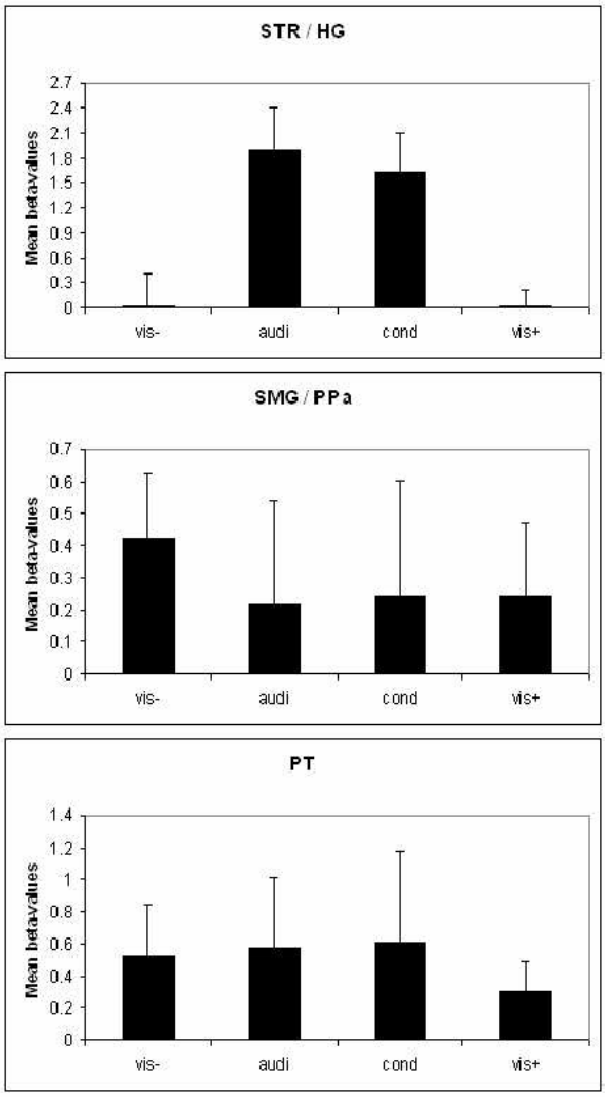


Fig. 4. Results of spherical ROI analysis. Mean t-values collected from three bilateral distinct regions in the temporo-parietal cortex (PT = planum temporale; STG (Heschl's gyrus) = superior temporal gyrus; SMG/PPa = supramarginal gyrus/planum parietale). Error bars refer to the standard deviation.

Table 5: Paired-sampled t-tests for three ROIs.

Region	*Comparison*	df	Sig. (2-tailed)
STR/HG	*vis- vs. audi*	14	.000
	vis- vs. cond	14	.000
	vis- vs. vis+	14	ns
	audi vs. cond	14	ns
	audi vs. vis+	14	.000
	cond vs. vis+	14	.000
SMG/PPa	*vis- vs. audi*	14	.044
	vis- vs. cond	14	.028
	vis- vs. vis+	14	.011
	audi vs. cond	14	ns
	audi vs. vis+	14	ns
	cond vs. vis+	14	ns
PT	*vis- vs. audi*	14	ns
	vis- vs. cond	14	ns
	vis- vs. vis+	14	.041
	audi vs. cond	14	ns
	audi vs. vis+	14	.051
	cond vs. vis+	14	.090

7.5 Discussion

The present study administered a multisensory association paradigm that was supposed to demonstrate that presentation of visual stimuli in isolation which had been paired with auditory stimuli activates regions associated with auditory perception. Interestingly, we noticed functional responses in the auditory association

cortex to visual stimuli during visual control (first run) and extinction (last run), that is before and after the paired presentation of flashes and sounds. As the finding of auditory activity exquisitely concurs with our predictions, the involvement of the right posterior Sylvian cortex compels demands a particular consideration. Accordingly, only one post-hoc explanation may account for this intriguing finding. As we applied an event related sparse temporal acquisition approach we surmise that participants instantly established an association between visual stimuli and scanner noise which consistently followed 3-5 s after presentation of flashes in each trial. Apparently, volunteers promptly learned to anticipate the onset of scanner noise each time they experienced visual stimuli followed by an auditory event. This learned association is reected by multisensory excitation of visual and auditory association cortex. This reasoning is exquisitely supported by the observation that auditory stimuli presented during the second and the third run activated the primary and secondary auditory cortex in the mid portion of the STG bilaterally while responses to visual stimuli in absence of auditory input clearly involved the bilateral SMG/PPa (more strongly in the first run) and the right PT more clearly the last run. The results delivered by the ROI analysis do not show a differential involvement of SMG/PPa and PT during the visual conditions. Even though the latter result is at odds with our predictions we would like to emphasize that at least the finding of hemodynamic responses in auditory association cortex triggered by multisensory bottom-up processing is in harmony with our hypotheses. Recent human and animal studies which have been investigating to what extent motor, visual, and somato-sensory stimuli induce responses in auditory regions report involvement of the same brain sites, namely the inferior parietal lobe (IPL), the posterior auditory association cortex, and the superior temporal sulcus (Foxe et al., 2002; Saito et al., 2005; Schroeder et al., 2001). Thus we will briefly refer to a selection of studies in more detail which have explored the functional roles of these regions in auditory perception. The PT has been coined a major 'computational hub' (Griffiths & Warren, 2002) and subserves a variety of genuine auditory functions, i.e. processing of auditory spectrotemporal information (Jancke et al., 2002; Meyer et al., 2005; Zaehle et al., 2004), temporal integration of sequential auditory events (Mustovic et al., 2003), neural presentation of pitch Xu et al. (2006), but also auditory imagery of linguistic and non-linguistic information (Bunzeck et al., 2005; Jancke & Shah, 2004; Pekkola et al., 2006). Neuroplastic changes in the PT have also been demonstrated by studies on comprehension of

sign language that observed that in congenitally deaf individuals the PT responds to visually presented linguistic information (MacSweeney et al., 2002; Petitto et al., 2000; Sadato et al., 2004, 2005). Interestingly a recent fMRI-study also identified the right posterior auditory cortex close to the PT as a cross-modal convergence zone evidently receiving auditory and somatosensory input (Foxe et al., 2002). In the context of the present study it is hence plausible to assume that right PT/STG activity which usually attests to auditory processing results from perceptually learned bottom-up associations between paired presentation of visual and auditory stimuli.

With respect to activation of the left V4 region (MTG/MTG) lighting up in the first run only we argue that decoding of the stimulus' colour led to its activity as this region has been attributed to colour perception (Schiller & Lee, 1991). We also obtained salient responses from the bilaterally SMG/PPa only during the first visual presentation. We assume that SMG/PPa activity in the visual habituation phase may also reect a visually triggered response of the auditory system. This region has been described as part of the macaque's dorsal auditory stream originating in the caudal part of the STG and projecting to the parietal cortex which preferentially responds to auditory spatial information (Rauschecker & Tian, 2000). Numerous functional imaging and clinical studies in humans support this evidence obtained from animal research as it has been demonstrated that the human IPL that accommodates the SMG/PPa is involved in associative auditory source localization (Adriani et al., 2003; Clarke et al., 2002; Weeks et al., 2000). Another function associated with this area has recently been described by (Gaab et al., 2003) who showed that SMG bilaterally (but mainly on the left) subserves working memory for tonal information and should therefore be considered region that is essential for higher auditory functions. A key role of the IPL and the adjacent parietal operculum evoked by auditory imagery of music has also recently been reported by an fMRI study which tested expressive and receptive aspects of cross-modal auditory-motor functions in professional pianists (Baumann et al., 2005). Responses to visual stimuli during the first runs also evoked responses in the left posterior STS which has been described as a heteromodal area which is presumed to bind information from unimodal sensory areas and thus help form crossmodal associations (Tanabe et al., 2005). According to a recent fMRI study the STS should be considered one region where auditory and visual information about objects is integrated (Beauchamp et al., 2004) and has also been noted to play a cardinal role in audiovisual speech perception (Ojanen et al., 2005; Wright et al.,

2003). Thus, we conjecture that in the context of the present study the STS might be involved in building a paired association between visual and auditory cues.

In accord with our predictions we discovered activation in the anterior peri-sylvian cortex, namely in the supramodal anterior insulae bilaterally in all runs and conditions of this study. This finding points to a general role of the anterior insulae in multisensory processing. Even though an immense number of neuroimaging studies have so far reported involvement of the anterior insulae related to a variety of sensory and cognitive tasks (Binder et al., 2004; Meyer et al., 2004; Plante et al., 2002; Wise et al., 1999; Wong et al., 2004) the precise function of this region is still unsettled. According to a recent review paper the anterior insulae play a vital role in visual-audio integration (Bamiou et al., 2003). As recently pointed out there is growing evidence which supports the view that the insula governs the detection of crossmodal coincidence (Calvert, 2001). In particular the right insula is supposed to support visual-auditory synchrony detection and thus plays a role in the synthesis of crossmodal cues on the basis of their temporal correspondence (Bushara et al., 1999). Based upon this converging evidence we reason that this region may be involved in integrating, establishing, and maintaining newly acquired crossmodal associations. Generally, this reasoning is in keeping with former imaging studies that tested classical defense conditioning and reported recruitment of bilateral insulae as a function of expectancy (Büchel et al., 1998; Hugdahl et al., 1995; Morris et al., 1998).

According to our starting point paired presentations of auditory and visual stimuli was meant to build up a tight multisensory relationship so that after a while the sole presentation of visual stimuli should sufice to elicit hemodynamic responses in auditory association cortex. Remarkably, we observed activation in auditory association cortex even during the first visual run prior to the paired presentation of auditory and visual cues. This intriguing finding becomes plausible once one considers the particular scanning protocol we used in the present study.

To avoid perceptual masking of auditory stimuli and physiological saturation of auditory cortices we applied a sparse temporal acquisition design which is made up of a consistent order of event-related stimulation in silence and subsequent scanner noise. Our study provides incidental evidence for the existence of multisensory devices binding together visual and auditory association cortices to instantly build up an internal representation of the tight relationship between red flashes and scanner

noise. However, the present data does not allow us to say unfailingly whether the auditory activation we observed during both the first and the last run emanated from conditioning or should be considered a reection of auditory imagery triggered by the anticipation of the scanner noise. Having said this we would like to emphasize the potential implications of this finding for current research on auditory perception. Sparse temporal protocols have been used successfully in the context of auditory fMRI (Hall et al., 1999; Schmidt et al., submitted) but the present finding strongly points to a presumptively fatal side-effect which researchers applying sparse temporal scanning should be aware of.

By all means, the current data evidently demonstrate that associative learning occurs quite automatically and involves cooperative interactions among many brain regions integrating specific sensory and less specific supramodal domains. Even though we are not able to say whether multisensory integration takes place early in the unisensory world or later at higher stages of processing, recently published data strongly indicates that visual input speeds up cortical processing of auditory signals at an early stage (Molholm et al., 2004; van Wassenhove et al., 2005). Bottom-up multisensory integration is more prevalent than previously recognized and confers a selective advantage in evolutionary terms. Thus, our data provide further buttressing to the observation that the integration of information from different sensory systems is a pivotal feature of perception (Ghazanfar & Schroeder, 2006) and is already occurring at or near the earliest processing stages (Schroeder et al., 2003). As recently outlined by Foxe et al. (2002, p. 543) "the early detection and localization of moving and perhaps threatening objects, has clear implications for survival and the presence of coincident sensory inputs is well known to improve detection and localization". In the context of our finding we reason that anticipation of canner noise promptly following the visual stimulus elicited responses which recruit multisensory (audio-visual) neural circuitries. In other words, we assume that the responses we observed in auditory association cortex in alliance with activation in heteromodal STS and insula should be considered part of a cross-modal visual-auditory network which converges as much sensory information as possible to enable rapid associative learning.

7.6 Conclusion

The results adduced by activation in auditory association cortex in the absence of auditory stimuli. In the present sparse temporal fMRI study we administered an associative learning paradigm which paired a visual stimulus (doubled flash) with an auditory stimulus (ringing of a telephone). Bilateral regions in the auditory association and heteromodal cortex (planum temporale, supramarginal gyrus) brought on a signal increase in the absence of auditory stimulation. This holds also to the run prior to the paired presentation as we noticed activity in association cortices which may reect an anticipatory response of multisensory mechanisms. Thus, we reason that the consistently paired presentation of visual stimuli followed by scanner noise sufficed to trigger auditory imaging reected by hemodynamic responses in nonvisual regions. Thus, our findings have implications for current research in the field of auditory fMRI as the anticipation of scanner noise in sparse temporal sampling studies may lead to conditioned response in auditory regions or trigger auditory imagery. Most importantly, our observation supports the present insights on the nature of densely intertwined multisensory networks during early stages of perceptual learning and also delivers novel findings in that it identifies sensory and higher brain regions involved in multi-sensory associative learning.

Acknowledgements

We are indebted to Conny Schmidt for assisting in MR data acquisition. Furthermore, we are grateful to Adam McNamara and three anonymous reviewers for providing helpful comments to improve the strength of the paper. During the preparation of this manuscript Martin Meyer and Simon Baumann were supported by Schweizer National Fonds (Swiss National Foundation) SNF 46234101, 'Short-term and long-term plasticity in the auditory system'.

7.7 References

Adriani M., Maeder P., Meuli R., Bellmann Thiran A., Frischknecht R., Villemure J.G., Mayer J., Annoni J.M., Bogousslavsky J., Fornari E., Thiran J.P., & Clarke S. (2003). Sound recognition and localization in man: specialized cortical networks and effects of acute circumscribed lesions. Brain Res. Exp. Brain Res., 153 , 591-604.

Annett M. (1970). A classification of hand preference by association analysis. Br. J. Psychol., 61 , 303-321.

Bamiou D.E., Musiek F.E., & Luxon L.M. (2003). The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res. Rev., 42 , 143-154.

Bandettini P., Jesmanowicz A., Van Kylen J., Birn R.M., & Hyde J.S. (1998). Functional MRI of brain activation induced by scanner acoustic noise. Magn. Reson. Med., 39 , 410-416.

Bao S., Chang E.F., Woods J., & Merzenich M. (2004). Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nature Neurosci., 7 , 974-981.

Baumann S., Koeneke S., Meyer M., Lutz K., & Jancke L. (2005). A network for sensorymotor integration. what happens in the auditory cortex during piano playing without acoustic feedback? Ann. N.Y. Acad. Scie., 1060 , 186-188.

Beauchamp M. S. Argall B.D., Bodurka J., Duyn J.H., & A. M. (2004). Unraveling multisensory integration: patchy organization within human sts multisensory cortex. Nature Neurosci., 7 , 1190-1192.

Belin P., Zatorre R.J., Hoge R., Evans A.C., & Pike B. (1999). Event-related fMRI of the auditory cortex. NeuroImage, 10 , 417-429.

Bernstein L.E., Auer E.T., Moore J.K., Ponton C.W., Don M., & Singh M. (2002). Visual speech perception without primary auditory cortex activation. NeuroReport, 13 , 311-315.

Binder J.R., Liebenthal E., Possing E.T., Medler D.A., & Ward D. (2004). Neural correlates of sensory and decision processes in auditory object identification. Nature Neurosci., 7 , 295-301.

Bosch V. (2000). Statistical analysis of multi-subject fMRI data: the assessment of focal activations. J. Magn. Reson. Imaging., 11 , 61-64.

Bunzeck N., Wuestenberg T., Lutz K., Heinze H.J., & Jancke L. (2005). Scanning silence: mental imagery of complex sounds. NeuroImage, 26 , 1119-1127.

Bushara K.O., Weeks R.A., Ishii K., Catalan M.J., Tian B., Rauschecker J., & Hallet M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nature Neurosci., 2 , 759-766.

Büchel C., & Dolan R. (2000). Classical fear conditioning in functional neuroimaging. Curr. Opin. Neurobiol., 10 , 219-223.

Büchel C., Morris J., Dolan R., & Friston K.J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron, 20 , 947-957.

Calvert G.A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb. Cortex , 11 , 1110-1123.

Calvert G.A., Bullmore E.T., Brammer M.J., Campbell R., Williams S.C.R., McGuire P.K., Woodruff P.W.R., Iversen S.D., & David A.S. (1997). Acitivation of auditory cortex during silent lipreading. Science, 276 , 593-596.

Calvert G.A., Campell R., & Brammer M.J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr. Biol., 10 , 649-657.

Clarke S., Bellmann Thiran A., Maeder P., Adriani M., Vernet O., Regli L., Cuisenaire O., & Thiran J.P. (2002). What and Where in human audition: selective deficits following focal hemispheric lesions. Brain Res. Exp. Brain Res., 147 , 8-15.

Edmister W.B., Talavage T.M., Ledden P.J., & Weisskopf R. (1999). Improved auditory cortex imaging using clustered volume acquisitions. Hum. Brain Mapp., 7 , 89-97.

Foxe J.J., Wylie G.R., Martinez A., Schroeder C.E., Javitt D.C., Guilfoyle D., Ritter W., & Murray M.M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. J. Neurophysiol., 88 , 540-543.

Friston K.J., Holmes A., Worsley K.J., Poline J., Frith C., & Frackowiak R. (1995). Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp., 2 ,189-210.

Gaab N., Gaser C., Zaehle T., Jancke L., & Schlaug G. (2003). Functional anatomy of pitch memory - an fMRI study with sparse temporal sampling. NeuroImage, 19 , 1417-1426.

Ghazanfar A.A., & Schroeder C.E. (2006). Is neocortex essentially multisensory? Trends Cog. Sc.i, 10 , 278-285.

Griffths T.D., & Warren J.D. (2002). The planum temporale as a computational hub. Trends Neurosci., 25 , 348-353.

Hall D.A., Haggard M.P., Akeroyd M.A., Palmer A.R., Summerfield A.Q., Elliott M.R., Gurney E.M., & Bowtell R.W. (1999). "Sparse" temporal sampling in auditory fMRI. Hum. Brain Mapp., 7 , 213-223.

Hugdahl K., Berardi A., Thompson W.L., Kosslyn S.M., Macy R., Baker D.P., Alpert N.M., & LeDoux J.E. (1995). Brain mechanisms in human classical conditioning: a PET blood ow study. NeuroReport, 6 , 1723-1728.

Jancke L. (1996). The hand performance test with a modified time limit instruction enables the examination of hand performance asymmetries in adults. Percept. Mot. Skills, 82 , 735-738.

Jancke L., Mirzazade S., & Shah N.J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in humansubjects. Neurosci. Lett., 266 , 125-128.

Jancke L., & Shah N. (2004). 'Hearing' syllables by 'seeing' visual stimuli. Eur. J. Neurosci., 19 , 2603-2608.

Jancke L.,Wüstenberg T., Scheich H., & Heinze H. (2002). Phonetic perception and the temporal cortex. NeuroImage, 15 , 733-746.

Komura Y., Tamura R., Uwano T., Nishjo H., & Ono T. (2005). Auditory thalamus integrates visual inputs into behavioral gains. Nature Neurosci., 8 , 1203-1209.

MacSweeney M., Amaro E., Calvert G.A., Campbell R., McGuire P.K., Williams S.C., Woll B., & Brammer M.J. (2000). Silent speechreading in the absence of scanner noise: an event-related fmri study. NeuroReport, 11 , 1729-1733.

MacSweeney M., Woll B., Campbell R., McGuire P.K., David A.S., Williams S.C.R., Suckling J., Calvert G.A., & Brammer M.J. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. Brain, 125 , 1583-1593.

McIntosh A.R., Cabeza R.E., & Lobaugh N.J. (1998). Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus. J. Neurophysiol., 80 , 2790-2796.

Meyer M., Steinhauer K., Alter K., Friederici A.D., & von Cramon D.Y. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang., 89 , 277-289.

Meyer M., Zaehle T., Gountouna E.V., Barron A., Jancke L., & Turk A. (2005). Spectrotemporal integration during speech perception involves left posterior auditory cortex. NeuroReport, 16 , 1985-1989.

Molchan S.E., Sunderland T., McIntosh A.R., Herscovitch P., & Schreurs B.G. (1994). A functional anatomical study of associative learning in humans. Proc. Natl. Acad. Sci. U. S. A., 91 , 8122-8126.

Molholm S., & Foxe J.J. (2005). Look 'hear', primary auditory cortex is active during lip-reading. NeuroReport, 16 , 123-124.

Molholm S., Ritter W., Javitt D.C., & Foxe J.J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. Cereb. Cortex , 14 , 452-465.

Morris J.S., Friston K.J., & Dolan R.J. (1998). Experience-dependent modulation of tonotopic neural responses in human auditory cortex. Proc. Biol. Sci., 265 , 649-657.

Moucha R., Pandya P.K., Engineer N.D., Rathbun D.L., & Kilgard M.P. (2005). Background sounds contribute to spectrotemporal plastcity in auditory cortex. Brain Res. Exp. Brain Res., 162 , 417-427.

Murray M.M., Foxe J.J., & Wylie G.R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. NeuroImage, 27 , 473-478.

Murray M.M., Michel C.M., Grave de Peralta R., Ortigue S., Brunet D., Andino S.G., & Schnider A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. NeuroImage, 21 , 125-135.

Mustovic H., Scheffer K., Di Salle F., Esposito F., Neuhoff J.G., Hennig J., & Seifritz E. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. NeuroImage, 20 , 429-434.

Nyberg L., Habib R., McIntosh A.R., & Tulving E. (2000). Reactivation of encoding-related brain activity during memory retrieval. Proc. Natl. Acad. Sci. U. S. A., 97 , 11120-11124.

Ojanen V., M ott onen R., Pekkola J., J a askel ainen I.P., Joensu R., Autti T., & Sams M. (2005).Processing of audiovisual speech in Broca's area. NeuroImage, 25 , 333-338.

Pekkola J., Ojanen V., Autti T., Jääskelainein I.P., M ot onen R., & Sams M. (2006). Attention to visual speech gestures enhances hemodynamic activity in the left planum temporale. Hum.Brain Mapp., 27 , 471-477.

Pekkola J., Ojanen V., Autti T., Jääskelainein I.P., M ot onen R., Tarkiainen W., & Sams M. (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3T. NeuroRe- port, 16 , 125-128.

Petitto L.A., Zatorre R.J., Gauna K., Nikelski E.J., Dostie D., & Evans A.C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: implications or the neural basis of human language. Proc. Natl. Acad. Sci. U. S. A., 97 , 13961-13966.

Plante E., Creusere M., & Sabin C. (2002). Dissociating sentential prosody from sentence processing: activation interacts with task demands. NeuroImage, 17 , 401-410.

Pruessmann K., Weiger M., Scheidegger M., & Boesiger P. (1999). SENSE: sensitivity encoding for fast mri. Magn. Reson. Med., 42 , 952-962.

Rademacher J., Morosan P., Schormann T., Schleicher A., Werner C., Freund H., & Zilles K. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. NeuroImage, 13 , 669-683.

Rauschecker J.P., & Tian B. (2000). Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. Proc. Natl. Acad. Sci. U. S. A., 97 , 11800-11806.

Sadato N., Okada T., Honda M., Matsuki K., Yoshida M., Kashikura K., Takei W., Sato T., Kochiyama T., & Yonekura Y. (2005). Cross-modal integration and plastic changes by lip movement,

random-dot motion and sign languages in the hearing and deaf. Cereb. Cortex , 15 , 1113-1122.

Sadato N., Yamada H., Okada T., Yoshida M., Hasegawa T., Matsuki K., Yonekura Y., & Itoh H. (2004). Age-dependent plasticity in the superior temporal sulcus in deaf humans: a functional mri study. BMC Neurosci., 15 , doi:10.1186/1471-2202-5-56.

Saito D.N., Yoshimura K., Kochiyama T., Okada T., Honda M., & Sadato N. (2005). Crossmodal binding and activated attentional networks during audio-visual speech integration: a functional MRI study. Cereb. Cortex , 15 , 1750-1760.

Schiller P.H., & Lee K. (1991). The role of the primate extrastriate area in V4 in vision. Science, 251 , 1251-1253.

Schmidt C.F., Meyer M., Baumann S., Baltes C., Boesiger P., & Jancke L. (submitted). Comparison of silent clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. NeuroImage.

Schroeder C.E., & Foxe J. (2002). Timing and laminar profile of converging inputs to multisensory areas of the macaque monkey. Brain Res. Cogn. Brain Res., 14 , 187-198.

Schroeder C.E., & Foxe J. (2005). Multisensory contributions to low-level, 'unisensory' processing. Curr. Opin. Neurobiol., 15 , 454-458.

Schroeder C.E., Lindsley R.W., Specht C., Marcovici A., Smiley J.F., & Javitt D.C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. J. Neurosci., 85 , 1322-1327.

Schroeder C.E., Smiley J., Fu K.G., McGinnis T., O'Connell M.N., & Hackett T.A. (2003). Anatomical mechanisms and functional implications of multisensory convergence in early corticalprocessing. Int. J. Psychophysiol., 50 , 5-17.

Tanabe H.C., Honda M., & Sadato N. (2005). Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. J. Neurosci., 25 , 6409-6418.

Tzourio-Mazoyer N., Landeau B., Papathanassiou D., Crivello F., Etard O., Delcroix N., Mazoyer B., & Joliot M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage, 15 , 273-289.

van Wassenhove V., Grant K.W., & Poeppel D. (2005). Visual speech speeds up the neural processing of auditory speech. Proc. Natl. Acad. Sci. U. S. A., 102 , 1181-1186.

Weeks R., Horwitz B., Aziz-Sultan A., Tian B., Wessinger C.M., Cohen L.G., Hallet M., & Rauschecker J.P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. J. Neurosci., 20 , 2664-2672.

Westbury C.F., Zatorre R.J., & Evans A.C. (1999). Quantifying variability in the planum temporale: a probability map. Cereb. Cortex , 9 , 392-405.

Wheeler M.E., Petersen S.E., & Buckner R.L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. U. S. A., 97 , 11125-11129.

Wik G., Elbert T., Fredrikson M., Hoke M., & Ross B. (1997). Magnetic brain imaging of extinction processes in human classical conditioning. NeuroReport, 8 , 1789-1792.

Wise R.J.S., Greene J., Buchel C., & Scott S.K. (1999). Brain regions involved in articulation. Lancet , 353 , 1057-1061.

Wong P.C.M., Parsons L.M., Martinez M., & Diehl R.L. (2004). The role of the insular cortex in pitch pattern perception: the effect of linguistics contexts. J. Neurosci., 24 , 9153-9160.

Wright T.M., Pelphrey K.A., Allsion T., McKeown M.J., & McCarthy G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. Cereb. Cortex , 13 , 1034-1043.

Xu Y., Gandour J., Talavage T., Wong D., Dzemidzic M., Tong Y., Li X., & Lowe M. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. Hum. Brain Mapp., 27 , 173-183.

Yoo S.S., Lee C.U., & Choi B.G. (2001). Human brain mapping of auditory imagery: eventrelated functional MRI study. NeuroReport, 12 , 3045-3049.

Zaehle T.,Wüstenberg T., Meyer M., & Jancke L. (2004). Evidence for rapid auditory perception as the foundation of speech processing - a sparse temporal sampling fMRI study. Eur. J.

8 Discussion, Conclusion & Outlook

This thesis consists of a combination of studies with the general aim to improve our knowledge on the extent and the mechanisms of plasticity in the auditory cortex. In order to address some of the still unresolved questions we started to set up several lines of experiments with various topics such as long- and short-term plasticity, plasticity of the connections from and to the auditory cortex etc. The studies included in this manuscript represent an excerpt from several of those lines. Because the individual studies were already thoroughly discussed in the light of the pertinent literature this section will mainly focus on issues concerning all or more than one of the studies. The discussion begins with a brief summary of the results and ends with a conclusion and a short outlook on future questions of interest.

8.1 A brief summary of the results

All of the four studies included in this thesis were undertaken to address the question whether (a) *the neural response of the auditory cortex to sound and* (b) *the strength of its connections to other modalities were modified by experience.* With the first two studies we showed that in contrast to previous reports musical experience does not specifically enhance the auditory response to tones derived from musical instruments but to tones in general. Furthermore we were able to demonstrate that the same enhancement is not based on transient attention, supporting the notion that musical experience modifies the structure of the auditory cortex. Nevertheless, according to data we derived from a modified task in which participants additionally were required to focus on music relevant sound features such as pitch and timbre, highly trained musicians differ from non-musicians also in early cortical attention processes. This result draws a picture of plasticity as a complex mixtures of sensory “bottom-up”- and prefrontal “top-down”-mechanisms. A side line of our analyses revealed cortical areas that specifically respond and presumably process spectral harmonics.

Bottom- up and top-down mechanisms also govern the establishment of crossmodal networks. That is one of the results of the third study in this thesis. We identified cortical areas being concurrently active during piano playing without acoustic feedback and piano listening and thus highlighted brain structures that play an important role in the information transfer between the auditory cortex and the motor cortex during piano performance. This audio-motor network which consisted mainly of secondary areas of the auditory, motor and somatosensory cortex showed stunning similarities to networks that are responsible for audio-motor control in speech. The fact that pianists compared to non-musician controls showed enhanced activity in the motor cortex during music listening demonstrates that musical experience does particularly strengthen audio-motor connections.

Finally, in an attempt to investigate short term plasticity of audio-visual connections, we applied a conditioning paradigm involving flashes and sounds. Our aim was to evoke activity in the auditory association cortex triggered by flashes which had previously been presented in combination with distinctive sounds. Surprisingly, we observed activity in the proposed areas, even before the audio-visual pairing was applied. Activation in the auditory cortex triggered by completely unrelated visual stimuli had not been demonstrated before. Potential explanations for this unreckoned finding include unintended fast audio-visual conditioning to scanner noise or innate audio-visual connections reacting to salient stimuli. The latter could part of a mechanism independent of the interpretation, these results underline our previous findings proposing a special role for the secondary auditory cortex in crossmodal information processing.

8.2 Are the extraordinary abilities of musicians acquired or innate?

In two of the studies included in this thesis we investigated long-term plasticity in the auditory cortex by comparing musicians to a control group. The extraordinary auditory training and the strong auditory skills of musicians are given facts described in the in the introduction. Provided that other factors such as age, handedness, education etc. are matched, we assumed that systematic differences in mental performance and brain activity between the groups are caused by the intensive music training and hence are indications of plasticity. In addition, it is assumed that all

innate or genetic factors are equally distributed among the groups. This assumption, however, is probably a simplification of the situation and it is only correct if everyone, independent of the genetic background, has the same chance to become a musician. This assumption might come close to truth dealing with musical amateurs. It is, however, obvious that this precondition is less given the higher the requirements for musicians are. The drop out fraction of those who can not meet the demanded requirements because they do not have the necessary genetic prerequisites for musical excellence is constantly rising on the way to the top for the internationally most recognized performers. In our studies, we mainly worked with students of the Zürich Conservatory. The extent to which the equal distribution of music relevant genetic factors at this level is skewed depends on the proportion of genetic and environmental factors predicting musical outcome. A number of studies were undertaken to investigate the influence of training and heredity on excellent music performance. An assessment of the amount of training clearly demonstrated a remarkably good correlation between time of practice and level of excellence in music (Ericsson et al., 1993; Ericson K.A., 1996). Biographical research, however, shows that musical talent in contrast to training is not very predictable (Howe et al., 1998; Sloboda and Howe, 1992; Sosniak, 1990). This indicates a less important role for heredity in becoming a successful musician. Therefore, intensive training is seen as the main factor for excellent performance. Nevertheless it is much too early to neglect innate factors for musical excellence. One way to address this question is to study the influence of music acquisition on cortical processes in early development. For example, in a recently completed a study (Meyer, Baumann et al. in progress), we investigated cortical activity during pre-attentive sound processing children of a Suzuki-Music-School and age-matched control children. Starting music training as early as 3 years of age is characteristic for children attending a Suzuki-School. The chance of having neurophysiological data from children at an early stage of music training and the collection of further data from participants at different stages of development allows us to investigate the progression of music-effects. This approach will help us to disentangle the effects of training and genetic factors and it will provide insight into the course of plasticity at different ages.

8.3 Expertise and increased neuronal representations of stimuli

One particular concept that was often mentioned in this manuscript but rarely closely defined, deserves further explanation. Pantev et al. (1998) brought the idea up that increased auditory evoked responses for instrumental tones compared to sine wave tones observed in musicians were an indication for an enlargement of cortical fields of neurons specifically sensitive to highly trained musical stimuli. This interpretation links alterations of neural responses due to training with data that demonstrated plasticity in animals. According to earlier studies in monkeys by Recanzone et al. (1993) the successful training of certain stimulus features (such as frequency) should result in larger neuronal representation of this feature. and therefore, following the interpretation of Pantev et al. (1998), it should lead to an enhanced MEG response. Because the proposed link between altered response amplitude and modified number of neuronal generators due to training is currently the only mechanistic explanation of the differences between musicians and nonmusicians in the auditory evoked responses, this concept is also relevant for the interpretation of the results of the first and the second study of this thesis. Although we basically agree with the idea that the altered response amplitude is an indication for modified neuronal response maps our data suggest a slightly more complex model than originally proposed.

The initial reports claiming enhanced auditory evoked responses in musicians compared to non-musicians (Pantev et al., 1998) showed stronger responses to instrumental tones compared to pure tones in musicians but no difference between the two tone classes in non-musicians. Based on the interpretation of increased electromagnetic responses as signs of plasticity due to training, it seems reasonable to expect stronger responses in musicians to trained instrumental tones than to untrained artificial pure tones. However, the first two studies included in this thesis showed that participants showed stronger cortical responses to instrumental tones regardless of musical expertise. Furthermore, direct comparisons between musicians and nonmusicians showed stronger responses in musicians for instrumental- as well as for pure-tones. Does this result mean that musicians react with stronger electrical responses and hence show larger neuronal representations for tones in general? Is it even an indication of an increased ability of musicians to perveive sounds?

The answer is probably more complex than a simple yes or no. Studies comparing hearing levels between musicians and controls show controversial results (Obeling

and Poulsen, 1999; Royster et al., 1991; Steurer et al., 1998). But even an increased hearing threshold of roughly 5 dB reported in one of those studies hardly explains the differences of the AEP between musicians and nonmusicians. The interpretation of the nature of the demonstrated neuronal representation enlargments might provide a clue to the problem. E.g., Ohl & Scheich (2005) are critical towards the interpretation by Recanzone et al. (1993) and argue that the representation of a single frequency by more neurons is unlikely to enhance the ability to discriminate the same frequency from others. In analogy, the simultaneous and homogenious representation of the sound of an instrument by more neurons would not enhance the processing of the same sound. This argument is reasonable if only one dimension (e.g. frequency) is considered and if the neuronal representation is only enlarged for a very specific point in this dimension (e.g. very specific frequency). Tones, however, are defined by a space of at least three dimensions: sound level, a temporal and spectral dimension. An enlargement of this three dimensional space involving more neurons would allow a higher resolution in all three dimensions. A larger representation of the auditory space in musicians could therefore facilitate the discrimination of specific sound features such as pitch, timbre etc. because these sound features defined by specific patterns of this space. In our experiments the auditory cortex responded to tones that spread over at least two (pure tones) or even over all three dimensions (complex instrumental tones). In this case, an enlarged auditory space providing a more fine-grained resolution and thus an enhanced perception of sound features in musicians would result in the activation of more neurons and therefore lead to stronger AEPs. So, an increased AEP to tones in musicians independent of whether these specific tones were trained or not, does not necessarily mean musicians feature a decreased hearing threshold compared to average but it might be an indication for being more sensitive to specific auditory aspects of tones due to an enhanced resolution for auditory features.

Such an interpretation of our results would have interesting implications beyond plasticity per se and it would represent an important finding for mechanisms summarized under the term “transfer-effects” that gain increasing interest. If an intensive training of specific stimuli, in our case musical tones, leads to an altered functional response to untrained stimuli, e.g. by enhancing the ability to resolve basic sound features common to auditory information, then the same training could also affect completely unrelated auditory abilities such as speech. The question whether a

successful musical training improves speech abilities is currently a topic of great interest in cognitive neuroscience (reviewed in Besson and Schon, 2001). Furthermore, the assumption of a transfer of training effects to untrained abilities as a general principle of our brain would have a considerable impact on the optimal way of learning in general. However, this interpretation requires confirmation e.g. by showing that musicians respond with an increased AEP to speech stimuli compared to nonmusicians before we postulate such transfer effects of musical training. Furthermore, extensive behavioral tests are required in order to increase the understanding of the extent to which musicians profit from an eventually enhanced auditory ability in the domain of language.

8.4 Top down and crossmodal influences on the auditory cortex

Reorganization of cortical maps as a result of sensory deprivation shown in animals (Merzenich et al., 1983; Rajan et al., 1993) but also as a result of a very intensive training (suggested for musicians by Pantev et al. (1998)) and by the data of the second study of this thesis, is certainly one of the more dramatic effects of neural plasticity. However, transient, top-down driven modulation of specific areas of our sensory cortex or use-dependent modification of preexisting associative connections between sensory and motor areas seem to be more adequate for fast adaptation to our environment or in case of short term learning (see also Singer, 1995). Due to the fact that research on the cerebral cortex is traditionally with a focus on a single modality, little is known about the mechanisms of sensory-motor or sensory-sensory interaction. The understanding of which cortical areas are crucial for such interactions is still vague. Whereas the discovery of mirror neurons (neurons in the motor related cortex that are active during the observation of actions) by Rizolatti et al. (1996) initiated a fast increase of research on visuo-motor connectivity and visuo-motor transformation mechanisms (reviewed in Burnod et al., 1999; Graziano et al., 1994), similar research on audio-motor interaction was almost absent at the start of this thesis (for an exception see Kohler et al., 2002). The data on audio-motor interaction during piano playing of musicians and non-musicians of the third study of this thesis provides information about cortical areas that play a particular role in such

interactions and in addition, it demonstrates the existence of plasticity of crossmodal or associative activity due to training. According to this study the secondary areas of the auditory, motor and somatosensory modalities such as the planum temporale (PT), the premotor cortex (PMC) and the so called secondary somatorsensory cortex (SII) are the most interesting candidate areas for audio-motor transformation processes. These findings which have already been published in a short version of the manuscript included in this thesis (Baumann et al., 2005) were recently replicated by similar audio-motor experiments with pianists (Bangert et al., 2006) but also by data derived from language-related audio-motor paradigms (Callan et al., 2006), from opera singers and actors (Kleber et al., Dick et al., presented at the Human Brain Mapping Conference 2006). Furthermore, a recent study by Chen et al. (in press) could demonstrate that activity in the PT and the PMC covaried while tapping with the hand to auditory presented rhythms. The variety of audio-motor tasks that seems to be dependent on the crossmodal transformation areas proposed by data of the third study of this thesis clearly supports our notion that we have observed an audio-motor network that is relevant far beyond piano playing. A comparison of our data with the data of various crossmodal experiments investigating all combinations of interactions between sensory and motor cortices even suggest that the secondary areas proposed above also play a similar role in the interplay with other modalities, such as the visual system. Whereas many experiments published following the discovery of mirror neurons highlighted the importance of the PMC in visuo-motor interactions, results from an increasing number of studies (including the fourth study of this thesis) provide congruent support for a particular role of the PT in audio-visual (Bunzeck et al., 2005; Jancke and Shah, 2004) and audio-somatosensory interactions (Foxe et al., 2002).

The data of the third study clearly prove that highlighted audio-motor interactions are susceptible to plasticity as a result of intensive piano training. Similarly, the results of the fourth study even suggest functional plasticity for audio-visual interactions on a time scale of minutes. Althought, further experiments are required to exclude that an unexpected association of the presented flashes with an unknown auditory percept has not existed in some of the participants before the data recording started. Nevertheless the results derived from the studies included in this thesis considerably increase the understanding of the interplay of the auditory cortex with other

modalities. Furthermore, several lines of evidence have been presented to illustrate that these interactions are susceptible to modification.

The fastest way to modify the functional properties of the auditory cortex in order to react on immediate changes in the environment is most likely provided by transient top down driven mechanisms such as attention. It is very unlikely that these temporary functional modifications alter the structural architecture of the cortex. It is rather expected that the electro-chemical properties of neurons is changed presumably involving modulatory circuits such as the cholinergic system (Sarter et al., 2005). The seminal study by Hillyard et al. (1973) showed for instance modulation of activity in the auditory cortex by shifting attention from one ear to the other. The results derived from the third study of this thesis demonstrate that in addition to activity in the auditory cortex driven by external stimuli, top down processes modulate crossmodal triggered activity. The activation in the auditory cortex during piano playing as well as premotor activation during music listening differed considerable in magnitude depending on whether activity was based on implicit or explicit processes. Likewise, the additional auditory potential component evoked by attention in highly trained musicians in contrast to nonmusicians that we describe in the second study clearly indicates that even modulating top-down processes are subject to long-term plasticity.

8.5 Conclusions & Outlook

The importance of neuronal plasticity for the understanding of learning and recovery processes in the brain has been recognized by neuroscientific research. Although several studies in the recent years indicated that plasticity occurs in the auditory cortex, the understanding of prerequisites and limits of plasticity is still sparse and many questions remain unanswered. According to the list at the beginning of this manuscript the thesis focuses on two questions (1) whether the activity in human auditory cortex is permanently modified by intensive auditory training and (2) whether the connectivity to other modalities is increased by specific training of crossmodal interactions. The first question was addressed by comparing the auditory evoked response to instrumental and sine wave tones in highly skilled musicians to

nonmusicians. The previous finding of increased responses of musicians to instrumental tones was confirmed and in addition a similar increase for sine wave tones was demonstrated. This result permits speculation on possible transfer effects of the trained ability beyond music to other fields such as language. Furthermore, our data excludes that the observed response increase in musicians compared to nonmusicians is based on transient attention which indicates that observed effect is permanent. Thus, permanent plasticity effects due to training are demonstrated which is affirming evidence to answer the first question. In addition, an increased auditory response to complex tones compared to sine wave tones in nonmusicians was used to identify cortical areas that play an important role in the processing of harmonics and timbre.

The identification of cortical areas that are involved in audio-motor and audio-visual interaction has been the basis to answer the second question. We revealed areas for each modality which presumably play a particular role in the information exchange between modalities. Several lines of evidence presented in this thesis suggest that the connections to the auditory cortex are susceptible to plastic modifications implying a affirmative answer to the second question.

Although the main objectives of this thesis were achieved, some questions remain unanswered and new questions have emerged. Additional experiments to address these questions are already in preparation or even in the process of publication. As mentioned in the discussion electrophysiological evidence for plasticity was observed in children attending a music school illustrating the developmental course of cortical modification due to training (Meyer, Baumann, Ringli & Jäncke, in preparation). A further logical step is certainly the systematic investigation of the range stimuli that evoke an enhanced response in musicians in order to define interesting candidate abilities that may profit from music training.

Further work is also in preparation to enhance the insight into the functioning of audio-motor interactions. Recently analyzed diffusion tensor imaging (DTI) data from musicians and nonmusicians suggests, based on calculated fiber tracks, direct connections between the proposed crossmodal transformation circuits in the secondary areas of the auditory cortex (PT, PP) and the motor cortex (PMC). We are currently investigating DTI measures with the aim to quantify the strengths of anatomical connections that would permit a comparison between musicians and nonmusicians. Furthermore we recorded EEG data from musicians and

nonmusicians to investigate frequency band coherence and phase locking. These data would permit the analysis of audio-motor integration with a time resolution of milliseconds, a time range that is relevant for fast feedback interactions. Furthermore, a close investigation of single crossmodal auditory and motor events by EEG event related potentials might have the potential to reveal the crossmodal influence of single keystrokes. Finally, understanding the precise circumstances and reasons that resulted in auditory activity when flashes were presented in the fourth study of this thesis clearly requires further experiments. This short outlook illustrates that the fields of auditory plasticity and crossmodal interaction are relatively new research areas and the potential for new, insightful experiments is immense. Therefore, this thesis is of particular relevance since it adds some of the first pieces to the complex puzzle of plasticity and interconnectivity of the auditory cortex.

8.6 References

Amunts, K., Schlaug, G., Jancke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., Zilles, K.,1997. Motor cortex and hand motor skills: Structural compliance in the human brain. Human Brain Mapping 5, 206-215

Bangert, M., Haeusler, U., Altenmuller, E.,2001. On practice: How the brain connects piano keys and piano sounds. Biological Foundations of Music 930, 425-428

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.J., Altenmuller, E.,2006. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. Neuroimage 30, 917-926

Baumann, S., Koeneke, S., Meyer, M., Lutz, K., Jancke, L.,2005. A network for sensory-motor integration - What happens in the auditory cortex during piano playing without acoustic feedback? Ann. N. Y. Acad. Sci. 1060, 186-188

Besson, M. and Schon, D.,2001. Comparison between language and music. Ann. N. Y. Acad. Sci. 930, 232-258

Bliss, T.V.P. and Gardnerm, A.R.,1973. Long-Lasting Potentiation of Synaptic Transmission in Dentate Area of Unanesthetized Rabbit Following Stimulation of Perforant Path. Journal of Physiology-London 232, 357-374

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H.J., Jancke, L.,2005. Scanning silence: Mental imagery of complex sounds. Neuroimage 26, 1119-1127

Burnod, Y., Baraduc, P., Battaglia-Mayer, A., Guigon, E., Koechlin, E., Ferraina, S., Lacquaniti, F., Caminiti, R.,1999. Parieto-frontal coding of reaching: an integrated framework. Experimental Brain Research 129, 325-346

Callan, D.E., Tsytsarev, V., Hanakawa, T., Callan, A.M., Katsuhara, M., Fukuyama, H., Turner, R.,2006. Song and speech: Brain regions involved with perception and covert production. Neuroimage 31, 1327-1342

Chen, C. and Tonegawa, S.,1997. Molecular genetic analysis of synaptic plasticity, activity-dependent neural development, learning, and memory in the mammalian brain. Annual Review of Neuroscience 20, 157-184

Ericsson, K.A., Krampe, R.T., Teschromer, C.,1993. The Role of Deliberate Practice in the Acquisition of Expert Performance. Psychological Review 100, 363-406

Ericsson, K.A., Ed. 1993. The Road of Expert Performance: Empirical Evidence from the Arts and Sciences, Sports and Games. Erlbaum. Mahwah, NJ.

Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., Goebel, R.,2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40, 859-869

Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W., Murray, M.M.,2002. Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. Journal of Neurophysiology 88, 540-543

Gaser, C. and Schlaug, G.,2003. Brain structures differ between musicians and non-musicians. Journal of Neuroscience 23, 9240-9245

Graziano, M.S.A., Yap, G.S., Gross, C.G.,1994. Coding of Visual Space by Premotor Neurons. Science 266, 1054-1057

Hebb, D.O.,1949. The Organization of Behavior. New York, Wiley

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W.,1973. Electrical Signs of Selective Attention in Human Brain. Science 182, 171-180

Howe, M.J.A., Davidson, J.W., Sloboda, J.A.,1998. Innate talents: Reality or myth? Behavioral and Brain Sciences 21, 399-+

Jancke, L. and Shah, N.J.,2004. 'Hearing' syllables by 'seeing' visual stimuli. European Journal of Neuroscience 19, 2603-2608

Jenkins, W.M., Merzenich, M.M., Ochs, M.T., Allard, T., Guicrobles, E.,1990. Functional Reorganization of Primary Somatosensory Cortex in Adult Owl Monkeys After Behaviorally Controlled Tactile Stimulation. Journal of Neurophysiology 63, 82-104

Kaas, J.H.,1983. What, If Anything, Is Si - Organization of 1St Somatosensory Area of Cortex. Physiological Reviews 63, 206-231

Kaas, J.H.,1991. Plasticity of Sensory and Motor Maps in Adult Mammals. Annual Review of Neuroscience 14, 137-167

Kaas, J.H. and Hackett, T.A.,1998. Subdivisions of auditory cortex and levels of processing in primates. Audiology and Neuro-Otology 3, 73-85

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., Rizzolatti, G.,2002. Hearing sounds, understanding actions: Action representation in mirror neurons. Science 297, 846-848

Kraus, N., Mcgee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B.,1996. Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science 273, 971-973

Kuriki, S., Kanda, S., Hirata, Y.,2006. Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. Journal of Neuroscience 26, 4046-4053

Merzenich, M.M., Jenkins, W.M., Johnston, P., Schreiner, C., Miller, S.L., Tallal, P.,1996. Temporal processing deficits of language-learning impaired children ameliorated by training. Science 271, 77-81

Merzenich, M.M., Kaas, J.H., Wall, J., Nelson, R.J., Sur, M., Felleman, D.,1983. Topographic Reorganization of Somatosensory Cortical Areas 3B and 1 in Adult Monkeys Following Restricted Deafferentation. Neuroscience 8, 33-55

Middlebrooks, J.C., Bierer, J.A., Snyder, R.L.,2005. Cochlear implants: the view from the brain. Current Opinion in Neurobiology 15, 488-493

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K.,2001. Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684-701

Obeling, L. and Poulsen, T.,1999. Hearing ability in Danish symphony orchestra musicians. Noise. Health 1, 43-49

Ohl, F.W. and Scheich, H.,2005. Learning-induced plasticity in animal and human auditory cortex. Current Opinion in Neurobiology 15, 470-477

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M.,1998. Increased auditory cortical representation in musicians. Nature 392, 811-814

Rajan, R., Irvine, D.R.F., Wise, L.Z., Heil, P.,1993. Effect of Unilateral Partial Cochlear Lesions in Adult Cats on the Representation of Lesioned and Unlesioned Cochleas in Primary Auditory-Cortex. Journal of Comparative Neurology 338, 17-49

Rauschecker, J.P.,1998. Parallel processing in the auditory cortex of primates. Audiology and Neuro-Otology 3, 86-103

Rauschecker, J.P., Tian, B., Hauser, M.,1995. Processing of Complex Sounds in the Macaque Nonprimary Auditory-Cortex. Science 268, 111-114

Rauschecker, J.P., Tian, B., Pons, T., Mishkin, M.,1997. Serial and parallel processing in rhesus monkey auditory cortex. Journal of Comparative Neurology 382, 89-103

Recanzone, G.H., Schreiner, C.E., Merzenich, M.M.,1993. Plasticity in the Frequency Representation of Primary Auditory-Cortex Following Discrimination-Training in Adult Owl Monkeys. Journal of Neuroscience 13, 87-103

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.,1996. Premotor cortex and the recognition of motor actions. Cognitive Brain Research 3, 131-141

Robertson, D. and Irvine, D.R.F.,1989. Plasticity of Frequency Organization in Auditory-Cortex of Guinea-Pigs with Partial Unilateral Deafness. Journal of Comparative Neurology 282, 456-471

Romanski, L.M., Bates, J.F., Goldman-Rakic, P.S.,1999. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. Journal of Comparative Neurology 403, 141-157

Royster, J.D., Royster, L.H., Killion, M.C.,1991. Sound Exposures and Hearing Thresholds of Symphony-Orchestra Musicians. Journal of the Acoustical Society of America 89, 2793-2803

Sarter, M., Hasselmo, M.E., Bruno, J.P., Givens, B.,2005. Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection. Brain Res. Brain Res. Rev. 48, 98-111

Schlaug, G., Jancke, L., Huang, Y.X., Staiger, J.F., Steinmetz, H.,1995. Increased Corpus-Callosum Size in Musicians. Neuropsychologia 33, 1047-&

Schneider, P., Scherg, M., Dosch, H.G., Specht, H.J., Gutschalk, A., Rupp, A.,2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nature Neuroscience 5, 688-694

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E.,2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. Journal of Neuroscience 23, 5545-5552

Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B.,2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport 16, 1781-1785

Sherman, S.M. and Spear, P.D.,1982. Organization of Visual Pathways in Normal and Visually Deprived Cats. Physiological Reviews 62, 738-855

Singer, W.,1995. Development and Plasticity of Cortical Processing Architectures. Science 270, 758-764

Sloboda, J.A. and Howe, M.J.A.,1992. Transitions in the Early Musical Careers of Able Young Musicians - Choosing Instruments and Teachers. Journal of Research in Music Education 40, 283-294

Sosniak, L.A.,1990. From Tyro to Virtuoso, A Long-Term Commitment to Learning. Piano Quarterly 29-&

Steurer, M., Simak, S., Denk, D.M., Kautzky, M.,1998. Does choir singing cause noise-induced hearing loss? Audiology 37, 38-51

Sterbing, S.J.,1996. Society of Neuroscience Abstracts 22, 889

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X.Q., Nagarajan, S.S., Schreiner, C., Jenkins, W.M., Merzenich, M.M.,1996. Language comprehension in language-learning impaired children improved with acoustically modified speech. Science 271, 81-84

Weinberger, N.M.,1993. Learning-induced changes of auditory receptive fields. Curr. Opin. Neurobiol. 3, 570-577

Acknowledgements

First I would like to thank Lutz Jäncke for giving me the opportunity to carry out my thesis in his lab and for keeping up such a great and motivating spirit in his group. I also wish to thank him for giving me a lot of freedom to test out my own ideas. I am sure this was a key element for the successful completion of my thesis.

I am also very grateful to Kevan Martin and Stephan Neuhauss for being my “responsible faculty member” and “thesis committee member” on behalf of the Faculty of Science, respectively. I particularly appreciate this because I am aware that both of them have to deal with a lot of requests for neuroscience Ph.D. theses which are conducted outside the faculty.

Throughout the years I highly appreciated Martin Meyer’s immense experience in our research field. He always supported me when problems turned up during my thesis. But I will also think back to those many times I could enjoy his great sarcastic humor when we simpy had a cup of tea in his office and were chatting about research questions.

Kai Lutz was an invaluable source for answers to questions concerning fMRI. But I could also discuss with him, often over a beer after work, all kinds of research problems and ideas for new projects. Some of them were successfully realized, many still remain ideas. Apart of that I have had a very joyful time with him in our common neat, little flat and during our many outdoor activities.

With Susan Koeneke I shared the many ups and downs of a Ph.D. student’s life during the three years of my thesis. We had a lot of fun in the lab, out and about in town or, as often as possible, on the snowboard. We cooperated successfully in several common studies and I learned to appreciate her organizing skills, her enthusiasm but also her critical reasoning when I occasionally got lost in too adventurous theories.

Of course, all my great colleagues contributed to the superb spirit and good cooperation within our group, so I would also like to thank Tino Zähle, Eveline Geiser, Petra Jäncke, Julia Blum, Mike Brügger, Elke Präg, Valentine Marcar and the former group members Michaela Esslen and Thomas Baumgartner for the wonderful time I had in Zurich. Given the bunch of fine new people who joined the lab shortly before I left, it is certain that you will keep up this spirit for the time to come.

A further, substantial source of motivation were the number of students I supervised during their master and project theses and also the ones who joined my seminars and courses in surprising numbers even on sunny summer days. They were very grateful, eager to learn new skills and committed to discuss actual questions in the field of neuroscience.

Writing up this thesis, I am very thankful to Susan Koeneke and Dave Hunter in Newcastle sorting out some of the worst language bits and pieces in parts of the thesis.

CURRICULUM VITAE

SIMON BAUMANN

PERSONAL DETAILS

Full name: Simon Baumann
Nationality: Swiss
Mother tongue: German

Date of birth: 21.01.1977
Place of birth: Affoltern, Switzerland
Place of origin: Grindelwald (BE), Switzerland

EDUCATION/QUALIFICATION

Sommer 2006- **Research Associate**
Research project: Development of auditory fMRI in monkeys
School of Neurology, Neurobiology & Psychiatry
Newcastle University

Fall 2003-
Sommer 2006 **Ph.D. thesis**
Research project: Plasticity and connectivity of the auditory cortex
Department of Neuropsychology, Institute for Psychology
University of Zurich
Supervisors: Prof. L. Jäncke, Prof. K. Martin (Institute of Neuroinformatics, University/ETH Zurich)
Participation in Ph.D. program of the Neuroscience Center Zurich

Fall 2002-
Spring 2003 **Diploma examinations (M.Sc.), Molecular Biology/Biochemistry**
Swiss Federal Institute of Technology (ETH) Zurich
Examinations: Biochemistry, Molecular Biology/Biophysics, Gentechnology, Neurosciences and Immunology.

Lectures in Cognitive Neurosciences and Imaging, University of Zurich

Fall 2001-
Summer 2002 **Diploma (Master) thesis**
Research project: The role of rapsyn in cytoskeletal anchoring of acetylcholine receptors.
Brain Research Institute, University of Zurich
Supervisor: Prof. C. Fuhrer, Prof. M. E. Schwab

Fall 1998-
2002 **Education in Molecular Biology/Biochemistry**
Swiss Federal Institute of Technology (ETH) Zurich

1998 **Matura**
Type C: Mathematics and Science
Kantonsschule Zug

PUBLICATIONS

Peer-Reviewed Publications

Baumann, S., Koeneke S., Meyer, M., Lutz, K., Jäncke, L. (2005) A network for sensory-motor integration: What happens in the auditory cortex during piano playing without acoustic feedback? *Ann. N. Y. Acad. Science.* 1060: 186-8.

Meyer, M., **Baumann, S.**, Jäncke, L. (2006) Electrical brain imaging evinces neural organisation of timbre perception. *NeuroImage,* 32 (4): 1510-23.

Jäncke, L., **Baumann, S.**, Koeneke, S., Meyer, M., Laeng, B., Peters, M., Lutz, K. (2006) Neural control of playing a reversed piano: Empirical evidence for an unusual cortical organization of musical functions. *Neuroreport.* 17(4), 447-51.

Meyer, M., **Baumann, S.**, Marchina, S., Jäncke, L. Audio-visual intertwining in the human brain during crossmodal associative learning. *BMC Neuroscience,* (in revision).

Schmidt, C., F., Meyer, M., **Baumann, S.**, Baltes, C., Boesiger, P., Jäncke, L. Comparison of a Clustered and Sparse Temporal Acquisition in fMRI during Tone and Speech Perception. *NeuroImage,* (in revision).

Baumann, S., Koeneke, S., Schmidt, C., F., Meyer, M., Lutz, K., Jäncke, L. A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Research,* (submitted).

Baumann, S., Meyer, M., Jäncke, L. Enhancement of Auditory Evoked Potentials in Musicians Reflects an Influence of Expertise but not Attention. *Journal of Cognitive Neuroscience, (submitted).*

Published Conference Abstracts (principal investigator)

Baumann, S., Meyer, M., Jäncke, L. Enhancement of Auditory Evoked Potentials in Musicians Reflects an Influence of Expertise but not Attention. *International Neuropsychological Society*. Zurich, Switzerland, July 26-30, 2006.

Baumann, S., Meyer, M., Jäncke, L. The neuroplastic enhancement of the auditory evoked potential in musicians is not an attention effect. *Human Brain Mapping, 12th Annual Meeting*. Florence, Italy, June 11-15, 2006.

Baumann, S., Koeneke, S., Koenig, H., Esslen, M., Lutz, K., Meyer, M., Jäncke, L. Electrophysiological evidence for an audio-motor integration network in pianists and non-musicians. *The 16th Meeting of the International Society for Brain Electromagnetic Topography.* Bern, Switzerland, October 5-8, 2005

Baumann, S., Koeneke, S., Schmidt, C., F., Meyer, M., Lutz, K., Jäncke, L. A network for audio-motor coordination in skilled pianists and non-musicians. *The Neurosciences and Music II – From Perception to performance*. Leipzig, Germany, May 5-8, 2005.

Baumann, S., Meyer, M., Jäncke, L. Time course of timbre perception – A comparison of complex and pure tones in LORETA. *Society for Neuroscience, 34th Annual Meeting.* San Diego, USA, October 23-27, 2004.

Baumann, S., Meyer, M., Schmidt, C. Lang, B., Peters, M. Jäncke, L. The case of a left-handed pianist playing a reversed keyboard: fMRI-findings from an auditory perspective. *2nd Vogt-Brodmann-Symposium on "The Convergence of Structure and Function"*. Jülich Research Centre, Jülich, Germany, April 2, 2004.

Talks

Baumann, S. How music training affects the auditory evoked potential. *CNRS Marseille*, Marseille, France, June 27th, 2006

Baumann, S. Plasticity in the auditory cortex – How professional music training affects the brain. *KLINEX – Colloquium for clinical and experimental Neuropsychology, University Hospital Zurich*. Zurich, Switzerland, April 12th, 2006

Baumann, S. Interaction of auditory and motor cortices in pianists and non-musicians. *Institute of Neuroscience, University of Newcastle.* Newcastle upon Tyne, UK, February 20th, 2006.

Baumann, S. Studying audio-motor integration in musicians. *Zurich fMRI Colloquium.* Zurich, Switzerland, March 22nd, 2005.

Baumann, S. Plasticity in the auditory cortex – An fMRI-study. *Neuroscience Centre Zurich Ph.D. Retreat*. Valens, Switzerland, May 14th, 2004.

Publications in preparation

Mapping of Mu- and Tau-rhythm during piano performance in musicians and controls.

Baumann, S., Koeneke S., Koenig H., Esslen M., Lutz, K., Meyer M., Jäncke, L (in preparation) Poster: *The 16th Meeting of the International Society for Brain Electromagnetic Topography.* Bern, Switzerland, October 5-8, 2005
Abstract published: (2005) *Brain Topography*, 18 (2): 125-166.

A short term pitch discrimination training investigated by sparsed fMRI

Baumann S., Meyer M., Jäncke L. (in preparation)

Connectivity between motor and auditory areas in pianists and controls revealed by DTI/fiber tracking.

Koeneke S., **Baumann S.**, Stämpfli P., Meyer M., Lutz K., Jäncke L. (in preparation)

EEG coherence and phase-looking shows evidence for audio-motor integration in pianists.

Koeneke S., **Baumann S.**, Koenig H., Esslen M., Meyer M., Lutz K., Jäncke L. (in preparation)

Children attending a Suzuki-Music-School show an increased "Mismatch Negativity" on instrumental tones compared to controls.

Meyer M., **Baumann S.**, Ringli M., Jäncke L. (in preparation)

Time course of auditory imagery illustrated by electroencephalographic imaging

Elmer S., Meyer M., **Baumann S.**, Jäncke L. (in preparation)